

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:38:19 ; Search time 53.1 Seconds  
(without alignments)  
9.677 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_154  
Perfect score: 15  
Sequence: 1 HSSKEKLRRRIKYC 15

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	46.7	327	1 DPPE_HAEIN	P45094 haemophilus
2	7	46.7	559	1 POTA_MYGE	P47288 mycoplasma
3	7	46.7	770	1 KLPA_EMENI	P28739 emericella
4	6	40.0	68	1 IPA3_YEAST	P01094 saccharomyc
5	6	40.0	90	1 HFQ_HAEIN	P44437 haemophilus
6	6	40.0	145	1 RRL7_ORYSA	Q9zst1 oryza sativ
7	6	40.0	159	1 GRE4_BUCAI	P57464 buchera ap
8	6	40.0	160	1 YMT0_YEAST	O4210 saccharomyc
9	6	40.0	177	1 PUR6_PYRHO	O38058 pyrococcus
10	6	40.0	181	1 NOHB_ECOLI	P31062 escherichia
11	6	40.0	181	1 TERS_LAMBD	P03707 bacterioph
12	6	40.0	189	1 NOHA_ECOLI	P31061 escherichia
13	6	40.0	242	1 YTXE_BACSU	P39064 bacillus su
14	6	40.0	318	1 YZ11_AQUAE	O66405 aquifex aeo
15	6	40.0	328	1 CEBB_CHICK	Q05826 gallus gall
16	6	40.0	334	1 DPPE_ECOLI	P37313 escherichia
17	6	40.0	372	1 BTOF_METJA	Q38694 methanococc
18	6	40.0	377	1 TDP_DROME	Q24318 drosophila
19	6	40.0	389	1 PYRD_DROME	P32748 drosophila
20	6	40.0	397	1 O85A_DROME	Q9vhs4 drosophila
21	6	40.0	509	1 G6PD_ANASP	P48992 anabaena sp
22	6	40.0	509	1 G6PD_NOSPO	P48848 nostoc punc
23	6	40.0	522	1 RECN_HELPJ	Q9zj80 helicobacte
24	6	40.0	524	1 RECN_HELPJ	O25943 helicobacte
25	6	40.0	526	1 BUTY_BOVIN	P18892 bos taurus
26	6	40.0	538	1 NRFA_HAEIN	P45017 haemophilus
27	6	40.0	552	1 Y4HP_RHISN	P50360 rhizobium s
28	6	40.0	556	1 PRIS_THIFE	P96095 thiobacillu
29	6	40.0	561	1 CCB2_DROME	Q9v8m2 drosophila
30	6	40.0	743	1 PMT5_YEAST	P52867 saccharomyc
31	6	40.0	795	1 SYFB_ECOLI	P07395 escherichia
32	6	40.0	875	1 UE3A_HUMAN	Q05086 homo sapien
33	6	40.0	885	1 UE3A_MOUSE	O08759 mus musculu

## RESULT 1

DPPE\_HAEIN  
ID DPPE\_HAEIN STANDARD; PRT; 327 AA.  
AC P45094;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE DIPEPTIDE TRANSPORT ATP-BINDING PROTEIN DPPE.  
GN DPPE OR H1184.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RD / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Sutton G., Fitchugh W., Fields C.A., Gocayne J.D.,  
RA McKenney K., Sutton G., Fitchugh W., Fields C.A., Kelley J.M.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelliey J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Georghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus  
influenzae Rd.";  
RL Science 269:496-512(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RD / KW20 / ATCC 51907;  
RX MEDLINE=96134971; PubMed=8550458;  
RA Preston A., Maskell D., Johnson A., Moxon E.R.;  
RA "Altered lipopolysaccharide characteristic of the I69 phenotype in  
Haemophilus influenzae results from mutations in a novel gene, isn.";  
RL J. Bacteriol. 178:396-402(1996).  
CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
FOR DIPEPTIDES. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO  
THE TRANSPORT SYSTEM (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
(ABC TRANSPORTERS).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
CC EMBL: U32798; AAC22837.1; -  
CC EMBL: U17295; AAA95975.1; -  
CC TIGR: H1184; -  
CC InterPro: IPR001617; -

## ALIGNMENTS

34 6 40.0 946 1 RHG4\_HUMAN  
35 6 40.0 1176 1 VPS8\_YEAST  
36 6 40.0 2210 1 RRPO\_TACV  
37 6 40.0 4036 1 RRPL\_DUGBV  
38 6 40.0 4969 1 RYNC\_RABIT  
39 5 33.3 26 1 PUTA\_KLEPN  
40 5 33.3 62 1 YM45\_CAEEL  
41 5 33.3 68 1 RL29\_PYRHO  
42 5 33.3 72 1 HST1\_ECOLI  
43 5 33.3 72 1 HST2\_ECOLI  
44 5 33.3 72 1 HST3\_ECOLI  
45 5 33.3 76 1 RS37\_YEAST

P98171 homo sapien  
P39702 saccharomyc  
P20430 tacaribe vi  
Q66431 dugbe virus  
P30957 oryctolagus  
P23725 klebsiella  
P34521 caenorhabdi  
O74094 pyrococcus  
P01559 escherichia  
O47185 escherichia  
P07965 escherichia  
P05759 saccharomyc

DR Pfam: PF00005; ABC\_tran; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 KW Peptide transport; Transport; Inner membrane; ATP-binding.  
 FT NP\_BIND 54 61 ATP (POTENTIAL).  
 SQ SEQUENCE 327 AA; 36917 MW; 0B50BDE197DA9BE CRC64;

Query Match 46.7%; Score 7; DB 1; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 2.7;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LRRERIK 13  
 |||||  
 Db 271 LRRERIK 277

RESULT 2  
 POT1\_MYCGE STANDARD; PRT; 559 AA.  
 ID P47288;  
 AC P47288;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE SPERMIDINE/PUTRESCINE TRANSPORT. ATP-BINDING PROTEIN POT1 HOMOLOG.  
 GN POT1 OR MG042.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmatataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann J.L., Buit C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritschman D.T., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
 RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Boff J.C., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium."  
 RL Science 270:397-403(1995).

CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
 CC FOR SPERMIDINE/PUTRESCINE. PROBABLY RESPONSIBLE FOR ENERGY  
 CC COUPLING TO THE TRANSPORT SYSTEM (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (POTENTIAL).  
 CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U39684; AAC71258.1; -  
 DR HSSP; P13569; INBD.  
 DR TIGR; MG042; -  
 DR InterPro; IPR001617; -  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 KW Transport; ATP-binding; Membrane.  
 FT NP\_BIND 40 47 ATP (POTENTIAL).  
 SQ SEQUENCE 559 AA; 65192 MW; DB8F04B26A110F6F CRC64;

Query Match 46.7%; Score 7; DB 1; Length 559;  
 Best Local Similarity 100.0%; Pred. No. 4.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKLRR 10  
 |||||

Db 173 KEKLRR 179

RESULT 3  
 KLPA\_EMENI STANDARD; PRT; 770 AA.  
 ID P28739;  
 AC P28739;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE KINESIN-LIKE PROTEIN KLPA.  
 GN KLPA.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; Emericella.  
 OX NCBI\_TaxID=5072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB20;  
 RX MEDLINE=93107178; PubMed=8416986;  
 RA O'Connell M.J., Meluh P.B., Rose M.D., Morris N.R.;  
 RT "Suppression of the bimC4 mitotic spindle defect by deletion of klpa,  
 RT a gene encoding a KAR3-related kinesin-like protein in Aspergillus  
 RT nidulans."  
 RL J. Cell Biol. 120:153-162(1993).  
 CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD  
 CC SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; X64603; CAA45887.1; -  
 DR PIR; A44337; A44337.  
 DR HSSP; P17119; 3KAR.  
 DR InterPro; IPR001752; -  
 DR Pfam; PF00225; kinesin; 1.  
 DR PRINTS; PR00380; KINESIN\_HEAVY.  
 DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 DR PROSITE; PS00671; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KW Motor protein; Microtubules; ATP-binding; Coiled coil.  
 FT DOMAIN 175 425 COILED COIL (POTENTIAL).  
 FT DOMAIN 426 770 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).  
 FT NP\_BIND 514 521 ATP (BY SIMILARITY).  
 SQ SEQUENCE 770 AA; 85800 MW; EFD0F0FF39B5C7EB CRC64;

Query Match 46.7%; Score 7; DB 1; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 5.9;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKLRR 10  
 |||||  
 Db 398 KEKLRR 404

RESULT 4  
 IPA3\_YEAST STANDARD; PRT; 68 AA.  
 ID P01094;  
 AC P01094;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE PROTEASE A INHIBITOR 3 (PROTEINASE INHIBITOR I(A)3).  
 GN IPA3 OR YMR174C OR YMR010.04C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;



```

[1]
RA  SEQUENCE FROM N.A.
RC  STRAIN=S288C;
RX  MEDLINE=91243884; PubMed=2037077;
RA  Schu P., Wolf D.H.;
RT  "The proteinase yscA-inhibitor, IA3, gene. Studies of cytoplasmic
RL  proteinase inhibitor deficiency on yeast physiology.";
RT  FEBS Lett. 283:78-84(1991).
[2]
RA  SEQUENCE FROM N.A.
RC  STRAIN=S288C / AB972;
RA  Churcher C.M., Barrrell B.G., Rajandream M.A., Walsh S.V.;
RL  Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
[3]
RA  SEQUENCE.
RA  Biedermann K., Montali U., Martin B., Svendsen I., Ottesen M.;
RT  "The amino acid sequence of proteinase A inhibitor 3 from baker's
RT  yeast.";
RL  Carlsberg Res. Commun. 45:225-235(1980).
CC  -!- FUNCTION: INHIBITOR FOR YEAST PROTEASE A (YSCA).
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL; X60050; CAA43650.1; -
DR  EMBL; Z49808; CAA89907.1; -
DR  PIR; A01334; IABY3.
DR  PIR; S16692; S16692.
DR  SGD; S0004786; PAI3.
KW  Protease inhibitor; Acetylation.
FT  MOD_RES 1 1 ACETYLATION.
FT  DOMAIN 2 23 NEARLY ALL OF THE INHIBITORY ACTIVITY IS
FT  PRESENT IN THIS REGION.
SQ  SEQUENCE 68 AA; 7707 MW; 8295BF7AF3E4A01A CRC64;

Query Match 40.0%; Score 6; DB 1; Length 68;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSKEKL 7
Db 14 SSKEKL 19
|||||

RESULT 5
ID HFQ_HAEIN STANDARD; PRT; 90 AA.
AC P44437;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HOST FACTOR-I PROTEIN (HF-I).
GN HFQ OR HI0411.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
[1]
RA  SEQUENCE FROM N.A.
RC  STRAIN=RD / KW20 / ATCC 51907;
RX  MEDLINE=95350630; PubMed=7542800;
RA  Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA  Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA  McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA  Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA  Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA  Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,

```

```

RA  Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagan N.S.M.,
RA  Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA  Venter J.C.;
RT  "Whole-genome random sequencing and assembly of Haemophilus
RT  influenzae Rd.";
RL  Science 269:496-512(1995).
CC  -!- FUNCTION: REQUIRED FOR PHAGE O BETA RNA-DIRECTED SYNTHESIS OF
CC  COMPLEMENTARY MINUS-STRAND RNA (BY SIMILARITY).
CC  -!- SUBUNIT: HOMOHXAMER (BY SIMILARITY).
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL; U32724; AAC22070.1; -
DR  TIGR; HI0411; -
KW  RNA-binding.
FT  INIT_MET 0 0 BY SIMILARITY.
SQ  SEQUENCE 90 AA; 10118 MW; 1C91D4ECF1F7F029 CRC64;

Query Match 40.0%; Score 6; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRERI 12
Db 14 LRRERI 19
|||||

RESULT 6
RR17_ORYSA
ID RR17_ORYSA STANDARD; PRT; 145 AA.
AC Q9ZST1;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 30S RIBOSOMAL PROTEIN S17, CHLOROPLAST PRECURSOR (CS17).
GN RPS17.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae;
OC Oryza
OX NCBI_TaxID=4530;
[1]
RA  SEQUENCE FROM N.A.
RC  STRAIN=CV. ILPOOM; TISSUE=Leaf;
RA  Lee J.-S., Eun M.-Y.;
RT  "Molecular cloning and characterization of three nuclear-encoded
RT  chloroplast precursor of ribosomal protein genes in rice.";
RL  Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC  -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC  -!- SIMILARITY: BELONGS TO THE S17P FAMILY OF RIBOSOMAL PROTEINS.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL; AF095707; AAC64969.1; -
DR  HSSP; P23828; 1RIP.
DR  InterPro: IPR000266; -
DR  Pfam: PF00366; Ribosomal_S17; 1.
DR  PRINTS: PR00973; RIBOSOMAL_S17.
KW  Ribosomal protein; Chloroplast; Transit peptide.
FT  TRANSIT 1 1 ? CHLOROPLAST.

```

```

FT CHAIN ? 145 30S RIBOSOMAL PROTEIN S17.
SQ SEQUENCE 145 AA; 15975 MW; A8D4693E5DC15E56 CRC64;

Query Match 40.0%; Score 6; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 RRERIK 13
Db 68 RRERIK 73

RESULT 7
ID GREABUCAI STANDARD; PRT; 159 AA.
AC P57464;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TRANSCRIPTION ELONGATION FACTOR GREAB (TRANSCRIPT CLEAVAGE FACTOR
DE GREAB)
GN GREAB OR BU384.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TOKYO 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -!- FUNCTION: NECESSARY FOR EFFICIENT RNA POLYMERASE TRANSCRIPTION
CC ELONGATION PAST TEMPLATE-ENCODED ARRESTING SITES. THE ARRESTING
CC SITES IN DNA HAVE THE PROPERTY OF TRAPPING A CERTAIN FRACTION
CC OF ELONGATING RNA POLYMERASES THAT PASS THROUGH, RESULTING IN
CC LOCKED TERNARY COMPLEXES. CLEAVAGE OF THE NASCENT TRANSCRIPT BY
CC CLEAVAGE FACTORS SUCH AS GREAB OR GREB ALLOWS THE RESUMPTION OF
CC ELONGATION FROM THE NEW 3' TERMINUS. GREAB RELEASES SEQUENCES OF
CC 2 TO 3 NUCLEOTIDES (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GREAB/GREB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001119; BAB13087.1; -.
CC PROSITE; PS00829; GREAB.1; 1.
CC DR PROSITE; PS00830; GREAB.2; 1.
CC KW Transcription regulation; DNA-binding; Coiled coil.
CC FT DOMAIN 10 27 COILED COIL (POTENTIAL).
CC SQ SEQUENCE 159 AA; 17983 MW; A400D7A95EFD1D1D3 CRC64;

Query Match 40.0%; Score 6; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 EKLRR 10
Db 13 EKLRR 18

RESULT 8
ID YMT0_YEAST STANDARD; PRT; 160 AA.

```

```

AC Q04210;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 19.2 KD PROTEIN IN SUB1-ARGRI INTERGENIC REGION.
GN YMR040W OR YMR532.05.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Odell C., Bowman S., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: TO YEAST YKL065C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z48502; CAA88406.1; -.
CC SGD; S0004643; YMR040W.
CC KW Hypothetical protein; Transmembrane.
CC FT TRANSMEM 3 23 POTENTIAL.
CC FT TRANSMEM 46 66 POTENTIAL.
CC FT TRANSMEM 104 124 POTENTIAL.
CC SQ SEQUENCE 160 AA; 19178 MW; D003EC7412559871 CRC64;

Query Match 40.0%; Score 6; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KEKLRR 9
Db 142 KEKLRR 147

RESULT 9
ID PUR6_PYRHO STANDARD; PRT; 177 AA.
AC O58058;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PHOSPHORIBOSYLAMINOIMIDAZOLE CARBOXYLASE CATALYTIC SUBUNIT
DE (EC 4.1.1.21) (AIR CARBOXYLASE) (AIRC).
GN PURE OR PH0320 OR PHCD015.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
CC -!- FUNCTION: THIS SUBUNIT CAN ALONE TRANSFORM AIR TO CAIR, BUT IN
CC ASSOCIATION WITH PURK, WHICH POSSESSES AN ATPASE ACTIVITY, AN
CC ENZYME COMPLEX IS PRODUCED WHICH IS CAPABLE OF CONVERTING AIR TO

```

CC CAIR EFFICIENTLY UNDER PHYSIOLOGICAL CONDITION (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: 1-(5-PHOSPHORIBOSYL)-5-AMINO-4-IMIDAZOLE-  
CC CARBOXYLATE = 1-(5-PHOSPHORIBOSYL)-5-AMINOIMIDAZOLE + CO(2).  
CC -!- PATHWAY: SIXTH STEP IN DE NOVO PURINE BIOSYNTHESIS.  
CC -!- SIMILARITY: TO OTHER BACTERIAL PURE, ALSO TO AIR CARBOXYLASE FROM  
CC FUNGI.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AP000002; BAA29394.1; -;  
DR InterPro: IPR000031; -;  
DR Pfam: PF00731; AIRC; 1.  
KW Purine biosynthesis; Lyase; Decarboxylase.  
SQ SEQUENCE 177 AA; 19439 MW; F86848BAA655703F CRC64;  
  
Query Match 40.0%; Score 6; DB 1; Length 177;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 4 KEKLR 9  
Db 145 KEKLR 150  
|||||  
  
RESULT 10  
ID NOHB\_ECOLI STANDARD; PRT; 181 AA.  
AC P31062; P75721;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE PROPHAGE QSR' DNA PACKAGING PROTEIN NU1 HOMOLOG.  
GN NOHB.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN [2]  
RP SEQUENCE OF 1-147 FROM N.A.  
RC STRAIN-K12 / W3110;  
RX MEDLINE=92334985; PubMed=1630906;  
RA Kotani H., Kawamura A., Takahashi A., Nakatsuji M., Hiraoka N.,  
RA Nakajima K., Takanami M.;  
RT "Site-specific dissection of E. coli chromosome by lambda terminase.";  
RL Nucleic Acids Res. 20:3357-3360(1992).  
CC -!- SIMILARITY: TO PHAGE LAMBDA DNA PACKAGING PROTEIN NU1.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF000161; AAC73661.1; -;

DR EMBL; D00928; -; NOT\_ANNOTATED\_CDS.  
DR EcoGene: EG11635; noHB.  
SQ SEQUENCE 181 AA; 20428 MW; E38D4FBFF174751C CRC64;  
  
Query Match 40.0%; Score 6; DB 1; Length 181;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 EKLRE 10  
Db 60 EKLRE 65  
|||||  
  
RESULT 11  
ID TERS\_LAMB STANDARD; PRT; 181 AA.  
AC P03707;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE TERMINASE SMALL SUBUNIT (DNA PACKAGING PROTEIN NU1).  
GN NU1.  
OS Bacteriophage lambda.  
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;  
OC Lambda phage group.  
OX NCBI\_TaxID=10710;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83189071; PubMed=6221115;  
RA Sanger F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B.;  
RT "Nucleotide sequence of bacteriophage lambda DNA.";  
RL J. Mol. Biol. 162:729-773(1982).  
RN [2]  
RP ATP-BINDING DOMAIN.  
RX MEDLINE=88172462; PubMed=2965248;  
RA Becker A., Gold M.;  
RT "Prediction of an ATP reactive center in the small subunit, gpNul, of  
RT the phage lambda terminase enzyme.";  
RL J. Mol. Biol. 199:219-222(1988).  
CC -!- FUNCTION: INVOLVED IN THE INITIATION OF LAMBDA DNA PACKAGING INFO  
CC THE PROHEAD. PROCESSES REPLICATING CONCATEMERIC DNA INTO PIECES  
CC OF UNIT LENGTH WITH COHESIVE ENDS.  
CC -!- SUBUNIT: HETEROOLIGOMER OF NU1 AND GPA.  
CC -!- SIMILARITY: TO BACTERIOPHAGE 21 TERMINASE SMALL SUBUNIT (GPI).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J02459; AAA96533.1; -;  
DR PIR; A04329; JVBPNL.  
KW DNA packaging; ATP-binding.  
FT NP\_BIND 31..36 ATP (BY SIMILARITY).  
SQ SEQUENCE 181 AA; 20441 MW; A25A5EAEF7BF18A4 CRC64;  
  
Query Match 40.0%; Score 6; DB 1; Length 181;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 EKLRE 10  
Db 60 EKLRE 65  
|||||  
  
RESULT 12  
ID NOHA\_ECOLI STANDARD; PRT; 189 AA.

AC P31061; P77152;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DE 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PROPHAGE QIN DNA PACKAGING PROTEIN NUI HOMOLOG.  
 GN NOHA.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RA "The complete genome sequence of Escherichia coli K-12.";  
 RT Science 277:1453-1474(1997).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=K12;  
 RC MEDLINE=97251357; PubMed=9097039;  
 RX Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,  
 RA Kasai H., Kashimoto K., Kimura S., Kitagawa M., Kitagawa M.,  
 RA Makino K., Mizobuchi K., Mori H., Mori T., Motomura K.,  
 RA Nakade S., Nakamura Y., Nishio Y., Oshima T., Saito N.,  
 RA Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,  
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;  
 RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 28.0-40.1 min region on the linkage map.";  
 RL DNA Res. 3:363-377(1996).  
 RN [3]  
 RP SEQUENCE OF 1-147 FROM N.A.  
 RC STRAIN=K12 / W3110;  
 RX MEDLINE=92334985; PubMed=1630906;  
 RA Kotani H., Kawamura A., Takahashi A., Nakatsuji M., Hiraoka N.,  
 RA Nakajima K., Takanami M.;  
 RA "Site-specific dissection of E. coli chromosome by lambda terminase.";  
 RT Nucleic Acids Res. 20:3357-3360(1992).  
 CC -!- SIMILARITY: TO PHAGE LAMBDA DNA PACKAGING PROTEIN NUI.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000252; AAC74621.1; -;  
 DR EMBL; D90798; BAA15252.1; -;  
 DR EMBL; D00927; -; NOT ANNOTATED\_CDS.  
 DR Ecogen; EG11634; nonA.  
 SQ SEQUENCE 189 AA; 21404 MW; FB5B1F844AB2C7C1 CRC64;  
 -----  
 Query Match 40.0%; Score 6; DB 1; Length 189;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 5 EKLRR 10  
 Db 60 EKLRR 65  
 RESULT 13  
 ID YTXE\_BACSU STANDARD; PRT; 242 AA.  
 AC P39064;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE HYPOTHETICAL 27.6 KDA PROTEIN IN ACUC 5'REGION (ORFB).  
 GN YTXE.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=95020526; PubMed=7934817;  
 RA Grundy F.J., Waters D.A., Takova T.Y., Henkin T.M.;  
 RA "Identification of genes involved in utilization of acetate and  
 RT acetoin in Bacillus subtilis.";  
 RL Mol. Microbiol. 10:259-271(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98048467; PubMed=9387221;  
 RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;  
 RA "Sequencing and functional annotation of the Bacillus subtilis genes  
 RT in the 200 kb rroB-dnaB region.";  
 RL Microbiology 143:3431-3441(1997).  
 CC -!- FUNCTION: MAY BE INVOLVED IN SOME TRANSPORT FUNCTION.  
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).  
 CC -!- SIMILARITY: BELONGS TO THE MOTB FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L17309; AAA68283.1; -;  
 DR EMBL; AF008220; AAC00301.1; -;  
 DR EMBL; Z99119; CAB14950.1; -;  
 DR PIR; S39642; S39642.  
 DR Subtilist; BG10366; YTXE.  
 DR InterPro; IPR001145; -;  
 DR Pfam; PF00691; OmpA; 1.  
 KW Hypothetical protein; Transport; Transmembrane.  
 FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 24 44 POTENTIAL.  
 FT DOMAIN 45 242 EXTRACELLULAR (POTENTIAL).  
 SQ SEQUENCE 242 AA; 27595 MW; 8BA7DDC103C1DFF5 CRC64;  
 -----  
 Query Match 40.0%; Score 6; DB 1; Length 242;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 6 KLRRR 11  
 Db 2 KLRRR 7  
 RESULT 14  
 ID YZ11\_AQUAE STANDARD; PRT; 318 AA.  
 AC O66405;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE HYPOTHETICAL PROTEIN AAL1.  
 GN AAL1.  
 OS Aquifex aeolicus.  
 OC Plasmid ecel.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 OX NCBI\_TaxID=63363;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;

```

RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- SIMILARITY: STRONG TO A.EOLICUS AA07 AND AA34.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000667; AAC07957.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 318 AA; 38423 MW; 4C2E5B005C0CAC7C CRC64;

Query Match 40.0%; Score 6; DB 1; Length 318;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 KLRRER 11
DB 140 KLRRER 145
|||||

RESULT 15
CEBB_CHICK
ID CEBB_CHICK STANDARD; PRT; 328 AA.
AC Q05826;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 01-FEB-1995 (Rel. 31, Last annotation update)
DE CCAAT/ENHANCER BINDING PROTEIN BETA (C/EBP BETA) (TRANSCRIPTION
DE FACTOR NF-M) (CCR PROTEIN).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93223673; PubMed=8467792;
RA Katz S., Kowenz-Leutz E., Mueller C., Meese K., Ness S.A.,
RA Leutz A.;
RT "The NF-M transcription factor is related to C/EBP beta and plays a
RT role in signal transduction, differentiation and leukemogenesis of
RT avian myelomonocytic cells.";
RL EMBO J. 12:1321-1332(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93259145; PubMed=84911193;
RA Burk O., Mink S., Ringwald M., Klemmner K.H.;
RT "Synergistic activation of the chicken mim-1 gene by v-myc and C/EBP
RT transcription factors.";
RL EMBO J. 12:2027-2038(1993).
CC -!- FUNCTION: HAS A ROLE IN SIGNAL TRANSDUCTION, DIFFERENTIATION AND
CC LEUKEMOGENESIS OF MYELOMONOCYTIC CELLS. BINDS TO THE MGF AND MIM-1
CC PROMOTERS AND ACTIVATES THE TRANSCRIPTION OF THESE GENES.
CC -!- SUBUNIT: BINDS DNA AS A DIMER.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN MYELOMONOCYTIC
CC CELLS.
CC -!- SIMILARITY: TO OTHER BZIP PROTEINS. STRONG, TO OTHER C/EBP
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z21646; CAA79760.1; -.
DR EMBL; X70813; CAA50144.1; -.
DR PIR; S35336; S35336.
DR TRANSFAC; T02022; -.
DR Pfam; PF00170; BZIP; 1.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
FT DNA_BIND 260 276
FT DOMAIN 289 317
FT LEUCINE-ZIPPER.
SQ SEQUENCE 328 AA; 35030 MW; 5AAE257F8213671C CRC64;

```

```

Query Match 40.0%; Score 6; DB 1; Length 328;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 6 KLRRER 11
DB 258 KLRRER 263
|||||

```

```

Search completed: September 15, 2001, 12:44:03
Job time: 344 sec

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:39:14 ; Search time 17.11 Seconds  
(without alignments)  
53.148 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_154  
Perfect score: 15  
Sequence: 1 HSSKEKLRRRIKYC 15

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 133343

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Listing first 45 summaries

Database : A.Geneseq.0601.\*

1: /cgnl\_9/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /cgnl\_9/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /cgnl\_9/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /cgnl\_9/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /cgnl\_9/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /cgnl\_9/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /cgnl\_9/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /cgnl\_9/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /cgnl\_9/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /cgnl\_9/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /cgnl\_9/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /cgnl\_9/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /cgnl\_9/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /cgnl\_9/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /cgnl\_9/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /cgnl\_9/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /cgnl\_9/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /cgnl\_9/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /cgnl\_9/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /cgnl\_9/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /cgnl\_9/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	14	93.3	15	21	AA199270
2	5	33.3	6	4	AA199295
3	5	33.3	6	15	AA199948
4	5	33.3	9	15	AA191019
5	4	26.7	4	20	AA193469
6	4	26.7	5	15	AA192114
7	4	26.7	5	17	AA1900407
8	4	26.7	5	19	AA191650
9	4	26.7	6	15	AA192104
10	4	26.7	6	16	AA1921203
11	4	26.7	6	20	AA1999137

12 4 26.7 6 21 AAY77109  
13 4 26.7 6 22 AAB36976  
14 4 26.7 7 12 AAR12307  
15 4 26.7 7 20 AAW99138  
16 4 26.7 7 21 AAY70721  
17 4 26.7 7 21 AAY77100  
18 4 26.7 7 21 AAY52428  
19 4 26.7 7 22 AAB50126  
20 4 26.7 8 8 AAF71271  
21 4 26.7 8 15 AAR61025  
22 4 26.7 8 21 AAB07782  
23 4 26.7 8 21 AAW90885  
24 4 26.7 9 14 AAR43772  
25 4 26.7 9 14 AAR43797  
26 4 26.7 9 14 AAR43815  
27 4 26.7 9 18 AAW39709  
28 4 26.7 9 18 AAW39520  
29 4 26.7 9 18 AAW54775  
30 4 26.7 9 19 AAW54799  
31 4 26.7 9 20 AAY53306  
32 4 26.7 9 20 AAY47943  
33 4 26.7 9 20 AAY40128  
34 4 26.7 9 20 AAY30908  
35 4 26.7 9 20 AAY26644  
36 4 26.7 9 21 AAY78499  
37 4 26.7 9 22 AAB46276  
38 4 26.7 9 22 AAB66548  
39 4 26.7 10 13 AAR28136  
40 4 26.7 10 14 AAR37218  
41 4 26.7 10 16 AAW21262  
42 4 26.7 10 16 AAR90154  
43 4 26.7 10 18 AAW41002  
44 4 26.7 10 18 AAW10108  
45 4 26.7 10 18 AAW18633

## ALIGNMENTS

RESULT 1  
RAY79270

ID AAY79270 standard; Peptide: 15 AA.

XX AC AAY79270;

XX DT 03-JUL-2000 (first entry)

XX DE PHELIIX peptide used to raise antibody.

XX KW PHELIIX; human; testis-specific; transcription factor;

XX KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;

XX KW therapy; diagnosis; vaccine; antibody.

XX OS Homo sapiens.

XX PN WO200012709-A2.

XX PD 09-MAR-2000.

XX PF 31-AUG-1999; 99WO-US20137.

XX PR 31-AUG-1998; 98US-0098610.

XX PR 31-OCT-1998; 98US-0106524.

XX PA (UROC-) UROGENESYS INC.

XX PA (AFAR/) AFAR D E.

XX PA (HUBE/) HUBERT R S.

XX PA (RAIT/) RAITANO A B.

XX PI Afar DE, Hubert RS, Raitano AB;

XX DR WPI; 2000-237872/20.

XX

PSA substrate pept  
Peptide used in ta  
ID1 plasmodium sur  
Peptide #2 cleavag  
Rat neuronal T-typ  
PSA substrate pept  
Cathpsin b-cleava  
Human brain T calc  
Sequence of new an  
Dynorphin-like pol  
Glutamic acid deca  
Human hyaluronid a  
MHC Class I allele  
MHC Class I allele  
MHC Class I allele  
HPV16 E6 peptide (  
HPV16 E6-encoded C  
Peptide from HPV 1  
Bcr-Abl epitope (a  
Immunogenic peptid  
Amino acid sequenc  
Human secreted pro  
Bcr-ABL-derived li  
SSX-4 derived pept  
HPV type 16 L1 pro  
Phage clone ns4 pi  
Cell-to-cell bindi  
IL-6 antagonist pe  
Hydroxymethylgluta  
Cys-bounded LALF-  
Substrate for TNFA  
Human agouti signa  
Peptide resulting

PT Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -

XX Example 5; Page 36; 62pp; English.

XX The present sequence is that of a peptide derived from human  
 CC PHELIIX (see AAY79269), a novel transcription factor that is  
 CC normally expressed only in testis tissue, but which is up-regulated  
 CC in prostate and some other cancers. The peptide was conjugated to  
 CC keyhole limpet haemocyanin and used to raise polyclonal antiserum  
 CC in rabbit. The antiserum demonstrated specificity for PHELIIX and  
 CC may therefore be useful for assessing the expression of PHELIIX in  
 CC patient samples.

XX Sequence 15 AA;

Query Match 93.3%; Score 14; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-08;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSSKEKLRRRIKY 14  
 DB | | | | | | | | | | | | | |

RESULT 2

AAP30295  
 ID AAP30295 standard; Protein; 6 AA.

XX AC AAP30295;

XX 20-APR-1992 (first entry)

XX Sequence of AAs 26-31 of the E. coli heat stable toxin  
 DE which determine an H-epitope.

XX Synthetic vaccine; antigen; allergen; immunological response;  
 KW antibody.

XX Escherichia coli.

XX EP93851-A.

XX 16-NOV-1983.

XX 11-MAR-1983; 83EP-0102392.

XX 15-MAR-1982; 82US-0358150.

PR 28-JAN-1983; 83US-0461802.

PR 12-JUN-1981; 81US-0272855.

PR 09-JAN-1981; 81US-0223558.

PR 16-DEC-1986; 86US-0942562.

XX (NYBL-) NEW YORK BLOOD CENT.

XX Hopp TP;

XX WPI; 1983-822049/47.

XX Synthetic vaccine - contains peptide residue coupled to higher  
 PT alkyl or alkenyl Gps. and with 6 amino acids in residue

XX Claim 50; Page 46; 54pp; English.

XX The inventors claim a synthetic vaccine which comprises a peptide  
 CC residue coupled to an alkyl or alkenyl gp. having at least 12C or  
 CC other lipophilic substance. The residue contains a sequence of 6 AAs  
 CC corresp. to the 50' of such AAs in a protein antigen or allergen where  
 CC the greatest local average hydrophilicity is found. Pref. the AAs in  
 CC the peptide do not exceed 50 residues, and they esp. contain 12-18  
 CC residues. The alkyl or alkenyl gp. pref. contains 12-24C and it is

CC pref. coupled to the terminal amino gp. of the residue opt. via a CO  
 CC gp. If a lipophilic substance is used, it is pref. palmitic, stearic,  
 CC behenic, oleic or mycolic acid.

XX Sequence 6 AA;

Query Match 33.3%; Score 5; DB 4; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKEK 6  
 DB | | | | |

RESULT 3

AAR59948  
 ID AAR59948 standard; peptide; 6 AA.

XX AC AAR59948;

XX 14-FEB-1995 (first entry)

XX Peptide signal sequence for treating E. coli infections.

XX Therapeutic; metabolic interactions; PSS; analogues.

XX Synthetic.

XX WO9416328-A.

XX 21-JUL-1994.

XX 30-DEC-1993; 93WO-US12679.

XX 30-DEC-1992; 92US-0997727.

XX (RATH/) RATH M.

XX Rath M;

XX WPI; 1994-249399/30.

XX Identifying peptide signal sequences in a protein - and use of  
 PT their synthetic analogues for treating or preventing, e.g.  
 PT cardiovascular and auto-immune disease, infections and cancer.

XX Claim 18; Page 10; 28pp; English.

XX The sequence is that of a peptide signal sequence which can be used  
 CC to treat E. coli infections.  
 CC See also AAR59944-83.

XX Sequence 6 AA;

Query Match 33.3%; Score 5; DB 15; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKEK 6  
 DB | | | | |

RESULT 4

AAR61019  
 ID AAR61019 standard; peptide; 9 AA.

XX AC AAR61019;

XX 23-APR-1995 (first entry)



XX Dynorphin-like polypeptide.  
DE Dynorphin; opioid analgesic; stable; stability; intravenous.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "MetYr"  
FT Misc-difference 8  
FT Modified-site 9 /note= "D-Glu"  
FT Modified-site 9 /note= "Arg-NH2"  
XX EP614913-A.  
PN 14-SEP-1994.  
PD 08-NOV-1985; 94EP-0107769.  
PF 09-NOV-1984; 84JP-0236076.  
PR (EISA ) EISAI CO LTD.  
XX Arakawa Y, Araki S, Ikeda M, Kaneko T, Nakazawa T;  
PI Tachibana S, Tsuchiya Y, Yamatsu K, Yoshino H;  
XX WPI; 1994-281157/35.  
XX New dynorphin polypeptide derivs. for use as analgesics - contg.  
PT a D-amino acid or an N-alkyl deriv. of a D- or L-amino acid for  
PT in vivo stability.  
XX Example 26; Page 27; 47pp; English.  
XX The patent discloses dynorphin-like polypeptides which are shorter  
CC than dynorphin (7 to 9 amino acids) and which contain either a D-  
CC amino acid and/or an N-alkyl deriv. of an L- or D-amino acid to  
CC provide in-vivo stability upon intravenous administration. The  
CC peptides have a much greater analgesic effect than dynorphin  
CC due to their greater stability in the blood (e.g. IC50 values can be  
CC as low as 0.04 nM when tested by the rabbit vas deferens method; cf.  
CC 17.4 nM for dynorphin).  
XX The present peptide is a specific example of the peptides.  
XX sequence 9 AA;  
SQ  
Query Match 33.3%; Score 5; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 LRRER 11  
Db 5 lrrer 9  
RESULT 5  
AAW83469  
ID AAW83469 standard; peptide; 4 AA.  
XX AAW83469;  
AC  
XX 08-MAR-1999 (first entry)  
DT Human growth hormone mutant peptide #29.  
XX Human; growth hormone; hGH; phagemid particle; enzyme substrate;  
KW fusion gene; phage protein coat.  
XX Homo sapiens.  
OS Synthetic.

XX US5846765-A.  
PN 08-DEC-1998.  
XX 16-MAY-1995; 95US-0441871.  
XX 03-DEC-1993; 93US-0161692.  
PR 03-DEC-1990; 90US-0621667.  
PR 10-APR-1991; 91US-0683400.  
PR 14-JUN-1991; 91US-0715300.  
PR 09-AUG-1991; 91US-0743614.  
PR 03-DEC-1991; 91WO-US09133.  
PR 06-APR-1992; 92US-0864452.  
PR 30-APR-1993; 93US-0050058.  
PR 05-APR-1995; 95US-0418928.  
PR 16-MAY-1995; 95US-0441871.  
XX (GETH ) GENENTECH INC.  
PA Matthews DJ, Wells JA, Zoller MJ;  
XX WPI; 1999-059058/05.  
XX Selection of polypeptide substrates - using phagemid particles  
PT displaying poly-peptide(s) as coat protein fusions  
PT Example 8; Column 32; 81pp; English.  
XX The present invention describes a method for selecting novel polypeptide  
CC substrates. The method comprises: (a) constructing a replicable  
CC expression vector comprising a transcription regulatory element operably  
CC linked to a gene fusion; (b) mutating the vector at one or more selected  
CC positions within the second gene thereby forming a family of related  
CC plasmids encoding substrate peptides; (c) transforming suitable host  
CC cells with the plasmids; (d) infecting the transformed host cells with a  
CC helper phage having a gene encoding the phage coat protein; (e) culturing  
CC the transformed infected host cells under conditions suitable for forming  
CC recombinant phagemid particles containing at least a portion of the  
CC plasmid and capable of transforming the host, the conditions being  
CC adjusted so that no more than a minor amount of phagemid particles  
CC display more than one copy of the fusion protein on the surface of the  
CC particle; (f) exposing the phagemid particles to a process capable of  
CC modifying at least one covalent bond of an amino acid in the substrate  
CC peptide of at least a portion of the phagemid particles; (g) contacting  
CC the family of exposed phagemid particles with an affinity molecule, where  
CC the affinity molecule has affinity for the amino acid residue having the  
CC modified covalent bond; and (h) separating the phagemid particles that  
CC bind to the affinity molecule from those that do not bind. The selection  
CC method is used for identifying enzyme substrates. The present sequence  
CC represents a human growth hormone mutant peptide from an example of the  
XX present invention.  
SQ Sequence 4 AA;  
Query Match 26.7%; Score 4; DB 20; Length 4;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 SSKE 5  
Db 1 sske 4  
RESULT 6  
AAR62114  
ID AAR62114 standard; peptide; 5 AA.  
XX AAR62114;  
AC  
XX 27-APR-1995 (first entry)  
DT  
XX

DE Hydrophilic motif from U1 snRNP 70K protein.  
 XX Small ribonucleoprotein complex; U1 snRNP: 70K protein; epitope;  
 KW autoantibody; immunoinfective cluster virus; nuclear protein antigen;  
 XX systemic rheumatic disorder; human immunodeficiency virus; HIV-1.  
 OS Homo sapiens.  
 XX  
 PN W09420141-A.  
 XX  
 PD 15-SEP-1994.  
 XX  
 PF 10-MAR-1994; 94WO-US02631.  
 XX  
 PR 11-MAR-1993; 93US-0029850.  
 XX  
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.  
 XX  
 PI Douvas A, Ehresmann G, Takehana Y;  
 XX  
 DR WPI; 1994-302689/37.  
 XX  
 PT Methods for treating immunoinfective cluster virus infections -  
 PT utilise antibodies or fragments characteristic of auto antibodies  
 PT produced by patients with rheumatic disorders  
 XX  
 PS Disclosure; Page 8; 106pp; English.  
 XX  
 CC The sequence RERR (AAR62113) is a preferred example of an alternating  
 CC acidic/basic amino acid, hydrophilic epitope motif, found in the  
 CC U1 snRNP 70K protein. It also occurs as RRERE and EREER (AAR62114  
 CC and AAR62115) in the 70K protein. The motif is also found in similar  
 CC form in immunoinfective cluster viruses. The motif serves as an  
 CC epitope for anti-viral antibodies and also for autoantibodies which  
 CC occur in high titre in patients suffering from systemic rheumatic  
 CC disorders. Sera from such patients could be used for treatment of  
 CC immunoinfective cluster virus (e.g. HIV,  
 CC EBV, rubella virus) infections.  
 XX  
 SQ Sequence 5 AA;  
 Query Match 26.7%; Score 4; DB 15; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 8 RRRER 11  
 DB 1 rrrer 4  
 RESULT 7  
 AAW00407  
 ID AAW00407 standard; peptide; 5 AA.  
 XX  
 AC AAW00407;  
 XX  
 DT 29-AUG-1996 (first entry)  
 XX  
 DE Interleukin-6 antagonist peptide.  
 XX  
 KW IL-6; antagonist; autoimmune disease.  
 XX  
 OS Synthetic.  
 XX  
 PN JP07324097-A.  
 XX  
 PD 12-DEC-1995.  
 XX  
 PF 30-MAY-1994; 94JP-0117259.  
 XX  
 PR 30-MAY-1994; 94JP-0117259.  
 XX

(DAIL ) DAICEL CHEM IND LTD.  
 (FUJI ) FUJISAWA PHARM CO LTD.  
 WPI; 1996-065476/07.  
 Interleukin 6 antagonist - useful for treating auto-immune diseases  
 Claims 2, 6; Pages 2, 3; 19pp; Japanese.  
 New IL-6 antagonists are provided which are of formula X-W-Y, in  
 which X is H or an amino-protecting group, Y is OH or a carboxy-  
 protecting group, and W is a peptide containing all or part of the  
 sequence as given in AAW00401, AAW00402, AAW00403 or AAW00404, where any  
 free mercapto groups in the sequence are optionally protected. The  
 present sequence is a specifically preferred partial sequence of AAW00401  
 and is itself claimed as a new chemical entity.  
 The IL-6 antagonists are useful for treating autoimmune diseases.  
 Sequence 5 AA;  
 Query Match 26.7%; Score 4; DB 17; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 SSKE 5  
 DB 2 sske 5  
 RESULT 8  
 AAW61650  
 ID AAW61650 standard; peptide; 5 AA.  
 XX  
 AC AAW61650;  
 XX  
 DT 27-OCT-1998 (first entry)  
 XX  
 DE Human reflex tears protein (cystatin S).  
 XX  
 KW Human; non-ocular disease; tear; cancer; breast; prostate.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09835229-A1.  
 XX  
 PD 13-AUG-1998.  
 XX  
 PF 06-FEB-1998; 98WO-AU00071.  
 XX  
 PR 07-FEB-1997; 97AU-0005009.  
 XX  
 PA (MACQ-) MACQUARIE RES LTD.  
 XX  
 PI (UNIX ) UNISEARCH LTD.  
 XX  
 PI Bolis S, Gooley AA, Herbert B, Molloy M, Morris C;  
 PI Walsh B, Willcox M, Williams KL;  
 XX  
 DR WPI; 1998-447373/38.  
 XX  
 PT Screening for non-ocular disease - by analysing tears for marker  
 PT proteins, particularly indicative of cancer and genetic disease,  
 PT also new proteins and nucleic acid encoding them  
 XX  
 PS Disclosure; Page 7; 14pp; English.  
 XX  
 CC The peptides AAW61650-W61652 were identified in human reflex tears,  
 CC in the method of the invention for screening for, or detecting,  
 CC non-ocular disease by analysing tears. Biochemicals, specifically  
 CC proteins, are isolated from tears, particularly by chromatography or  
 CC electrophoresis, especially two-dimensional polyacrylamide gel  
 CC electrophoresis (2D-PAGE), then detected, e.g. with labelled specific  
 CC reagents, in (radio)immunoassay. The method is used to detect cancer,

CC particularly of breast or prostate, or a genetic disease, in humans or  
 CC animals.

SQ Sequence 5 AA;

Query Match 26.7%; Score 4; DB 19; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKE 5  
 ||||  
 Db 2 sske 5

RESULT 9

AA62104  
 ID AAR62104 standard; peptide; 6 AA.

XX AC AAR62104;

XX DT 27-APR-1995 (first entry)

XX DE Hydrophilic motif from nuclear protein antigens.

XX KW Small ribonucleoprotein complex; U1 snRNP; 70K protein; epitope;  
 KW autoantibody; immunoinfective cluster virus; nuclear protein antigen;  
 KW systemic rheumatic disorder; human immunodeficiency virus; HIV-1;  
 KW centromere CENP-B; thyroglobulin-h; thyroid peroxidase; scleroderma;  
 KW systemic lupus erythematosus.

XX OS Homo sapiens.

XX PN WO9420141-A.

XX PD 15-SEP-1994.

XX PF 10-MAR-1994; 94WO-US02631.

XX PR 11-MAR-1993; 93US-0029850.

XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX PI Douvas A, Ehresmann G, Takehana Y;

XX DR WPI; 1994-302689/37.

XX PT Methods for treating immunoinfective cluster virus infections -  
 PT utilise antibodies or fragments characteristic of auto antibodies  
 PT produced by patients with rheumatic disorders

XX PS Disclosure; Page 8; 106pp; English.

XX CC This sequence is an example of an alternating acidic/basic amino  
 CC acid, hydrophilic motif possibly found in nuclear protein antigens.  
 CC As well as occurring in normal human proteins, the motif is found  
 CC in similar form in immunoinfective cluster viruses. The motif  
 CC serves as an epitope for anti-viral antibodies and also for  
 CC autoantibodies which occur in high titre in patients suffering  
 CC from systemic rheumatic disorders. Sera from such patients could  
 CC be used for treatment of immunoinfective cluster virus (e.g. HIV,  
 CC EBV, rubella virus) infections.

XX SQ Sequence 6 AA;

Query Match 26.7%; Score 4; DB 15; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RRER 11

Db 1 rrer 4

RESULT 10

AAW21203

XX ID AAW21203 standard; peptide; 6 AA.

XX AC AAW21203;

XX DT 29-JUL-1997 (first entry)

XX DE Farnesyl synthetase derived signal oligopeptide #3.

XX KW Hydrophilic; signal oligopeptide; hydrophilicity maxima; vaccine; SIV;  
 KW competitive inhibitor; feedback regulator; synthesis; gastrin precursor;  
 KW charge; polarity; farnesyl synthetase; plasminogen activator inhibitor 1;  
 KW hydroxymethylglutaryl coenzyme A reductase; glucagon precursor; rhesus;  
 KW gonadoliberin precursor; plasminogen activator inhibitor 2; prorenin;  
 KW Alzheimer amyloid A4; corticotropin releasing factor binding protein;  
 KW apolipoprotein E; herpes virus 1 glycoprotein B; HSV1; human; OMVVS;  
 KW herpes virus 2 glycoprotein B; HSV2; collagenase; apolipoprotein A;  
 KW treponema pallidum membrane protein; TMPA; islet amyloid polypeptide;  
 KW fibroblast MMP1; schistosoma elastase precursor; schistosomin;  
 KW hepatitis delta antigen; rev protein; HIV; VILV; angiotensinogen.

XX OS Homo sapiens.

XX PN WO9519568-A1.

XX PD 20-JUL-1995.

XX PF 12-JAN-1995; 95WO-US00575.

XX PR 14-JAN-1994; 94US-0182248.

XX PA (RATH//) RATH M.

XX PI Rath M;

XX DR WPI; 1995-263953/34.

XX PT Identifying signal oligopeptide(s) in protein sequence(s) - shown as  
 PT regions of max. hydrophilicity, used in modulating communication  
 PT between protein(s)

XX PS Claim 5; Page 23; 88pp; English.

XX CC The sequences given in AAW21201-560 represent hydrophilic signal oligo-  
 CC peptides. These signal oligopeptides are localised on the surface  
 CC of the protein and are represented by the hydrophilicity maxima of  
 CC the protein. These peptides are enriched in charged amino acids  
 CC arranged with neutral spacer amino acids. The specific signal  
 CC character of these oligopeptides is determined by a characteristic  
 CC combination of conformation and charge within the signal sequence.  
 CC These oligopeptides may be used as vaccines in the treatment of  
 CC human disease, as competitive inhibitors to prevent or reduce the  
 CC metabolic action or interaction of a selected protein by blocking  
 CC its specific signal sequences, or as therapeutic agents to function  
 CC as feedback regulators to reduce synthesis rate of a selected protein.  
 CC These peptides may be modified by omitting one or more amino acids at  
 CC the N- and/or C-terminal, by substituting one or more amino acids  
 CC without consideration of charge and polarity, by substituting one or  
 CC more amino acids with amino acid residues with similar charge and/or  
 CC polarity, by omitting one or more amino acids or a combination of these.

XX SQ Sequence 6 AA;

Query Match 26.7%; Score 4; DB 16; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RRER 11

||||

Db 3 rrer 6

RESULT 11  
AAW99137  
ID AAW99137 standard; peptide; 6 AA.  
XX AC  
XX AAW99137;  
DT 17-MAY-1999 (first entry)  
XX  
XX  
XX Peptide #1 cleavage site specific for PSA-proteolytic enzyme.  
DE  
XX  
XX Cleavage site; prostate specific antigen; PSA; proteolytic enzyme;  
KW prodrug; prostatic glandular cell; sesquiterpene-gamma-lactone;  
KW cell proliferation; prostate cancer; breast cancer.  
XX  
XX Synthetic.  
XX W09852966-A1.  
XX  
XX 26-NOV-1998.  
XX  
XX PF 19-MAY-1998; 98WO-US10285.  
XX  
XX 30-MAR-1998; 98US-0080046.  
PR 19-MAY-1997; 97US-0047070.  
XX  
XX (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
PA  
XX Christensen SB, Denmeade SR, Issacs JT, Lilja H;  
PI WPI; 1999-131688/11.  
XX  
XX Peptides used in compositions to treat prostate-specific  
PT antigen-producing cell proliferation - have amino-acid sequence with  
PT cleavage site specific for PSA-proteolytic enzyme  
XX  
XX Claim 8; Page 44; 58pp; English.  
XX  
XX The present sequence represents a peptide comprising an amino-acid  
CC sequence with a cleavage site specific for an enzyme with the  
CC proteolytic activity of prostate-specific antigen (PSA). The peptide  
CC can be used in compositions for the treatment of PSA-producing cell  
CC proliferative disorders, both benign and malignant, such as prostate  
CC breast cancers, to detect PSA-producing tissue, to select PSA-  
CC activatable prodrugs substantially specific for target tissue producing  
CC PSA-producing cells to determine activity of PSA in samples and to image  
CC PSA-producing tissue. The peptide can be linked to therapeutic drugs,  
CC such as thapsigargin, estafiatin, grossheimin, inuchineolide, arglavin  
CC and derivatives such as thapsigargin, to form prodrugs thereby  
CC substantially inhibiting non-specific toxicity by activating drug  
CC locally at PSA-producing when enzymatic cleavage takes place.  
XX  
SQ Sequence 6 AA;

Query Match 26.7%; Score 4; DB 20; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSSK 4  
Db 1 hssk 4  
|||||

RESULT 12  
AAW77109  
ID AAW77109 standard; peptide; 6 AA.  
XX AC  
XX AAW77109;  
XX  
XX 08-MAY-2000 (first entry)  
DT

PSA substrate peptide linker SEQ ID NO:13.  
Prostate-specific antigen; PSA substrate peptide; drug complex;  
targeted delivery; cytotoxic drug; anticancer; prostate cancer;  
androgen-independent; proteolysis.  
Synthetic.  
W0200001419-A1.  
13-JAN-2000.  
06-JUL-1999; 99WO-US15126.  
06-JUL-1998; 98US-0110822.  
(JCRT-) JCRT RADIATION ONCOLOGY SUPPORT SERVICES.  
(UYNE-) UNIV NORTHEASTERN.  
(BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
D'Amico AV, Bubley GJ, Jebaratnam DJ, Weinberg JS;  
WPI; 2000-160730/14.  
New drug complexes comprising a targeting carrier molecule and a  
cytotoxic drug linked by a peptide which is cleaved by molecules near  
target cells, used particularly for treating prostate cancer -  
Claim 2; Page 7; 35pp; English.  
The invention relates to a novel drug complex comprising a targeting  
carrier molecule which, when introduced into an individual, is  
selectively distributed to prostate tissue, bone or both; a peptide  
which is a substrate for prostate specific antigen (PSA); and a cytotoxic  
drug which is toxic to prostate cancer cells, particularly androgen-  
independent cancer cells. The peptide links the targeting molecule and  
the cytotoxic drug. PSA is a protease that is present in the  
microenvironment of malignant prostatic epithelium. Cleavage of the  
peptide linker by PSA liberates the cytotoxic drug at the site where it  
is required. The drug complexes can be used to deliver agents to  
specific cells, particularly prostate cancer cells. The drug complex of  
the invention provides a prostate cancer treatment which has higher  
efficacy and lower toxicity than presently available treatments.  
Sequences AAW7097-Y77111 represent peptides which are substrates for PSA  
and which can be used as linkers in drug complexes of the present  
invention.  
Sequence 6 AA;

Query Match 26.7%; Score 4; DB 21; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSSK 4  
Db 1 hssk 4  
|||||

RESULT 13  
AAB36976  
ID AAB36976 standard; peptide; 6 AA.  
XX AC  
XX AAB36976;  
XX  
XX 28-FEB-2001 (first entry)  
DT  
XX Peptide used in targeting prostate tissue.  
DE  
XX Prostate; cancer.  
KW  
XX Synthetic.  
OS

XX WO200066175-A2.  
 XX 09-NOV-2000.  
 XX 27-APR-2000; 2000WO-US11542.  
 XX 30-APR-1999; 99US-0131809.  
 XX (SLIL-) SLIL BIOMEDICAL CORP.  
 XX Frydman B, Marton LJ;  
 XX WPI; 2001-031721/04.  
 XX Compositions comprising a peptide substrate conjugated to a cytostatic  
 PT or cytotoxic agent, cleaved by enzymes, useful for treating  
 PT prostatitis, benign prostatic hyperplasia, prostate cancer -  
 XX Claim 25; Page 35; 194pp; English.  
 XX The present invention relates to a composition containing a peptide  
 CC substrate conjugated to a cytostatic or cytotoxic agent, cleaved by  
 CC enzymes. The composition can be used to treat prostate disease, such as  
 CC prostate cancer.  
 XX SQ Sequence 6 AA;  
 Query Match 26.7%; Score 4; DB 22; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 HSSK 4  
 Db 1 hssk 4  
 RESULT 14  
 ID AAR12307 standard; Protein; 7 AA.  
 XX AAR12307;  
 AC AAR12307;  
 XX 29-AUG-1991 (first entry)  
 DT DT  
 XX ID1 plasmodium surface peptide.  
 DE DE  
 XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria;  
 KW hybrid.  
 XX Plasmodium falciparum.  
 OS OS  
 XX EP432965-A.  
 PN PN  
 XX 19-JUN-1991.  
 PD PD  
 XX 06-DEC-1990; 90EP-0313257.  
 PF PF  
 XX 08-DEC-1989; 89US-0447746.  
 PR PR  
 XX (SMIK ) SMITHKLINE BEECHAM.  
 PA (USSA ) US SEC OF THE ARMY.  
 XX (BIOM-) BIOMEDICAL RES INST.  
 XX Gross MS, Gordon DM, Hollingdale MR;  
 PI WPI; 1991-179771/25.  
 XX Polypeptide comprising immunogenic determinants from P falciparum  
 PT - for vaccine against malaria infection in humans.  
 XX Claim 1; Page 16; 18pp; English.

XX The peptide is a sporozoite neutralising epitope from the 1st flank-  
 CC ing region of the CS protein of plasmodium. It can be used in a  
 CC vaccine for protection against malaria. The peptide is pref. linked  
 CC to a 2nd peptide from the 2nd flanking domain and the resulting  
 CC polypeptide fused to a carrier protein, e.g tetanus toxoid, dip-  
 CC theria toxin or cholera B toxin. A preferred vaccine comprises 81  
 CC N-terminal AAs of the influenza virus nonstructural protein 1  
 CC (NS181), fused, via a synthetic linker, to a 1st flanking region of  
 CC the CS protein, which is itself fused to a 2nd flanking region of  
 CC the CS protein.  
 CC See also AAR12306-R12311 and AAR13175-R13179.  
 XX SQ Sequence 7 AA;  
 Query Match 26.7%; Score 4; DB 12; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 EKLR 8  
 Db 1 eklr 4  
 RESULT 15  
 ID AAW99138 standard; peptide; 7 AA.  
 XX AAW99138;  
 AC AAW99138;  
 XX 17-MAY-1999 (first entry)  
 DT DT  
 XX Peptide #2 cleavage site specific for PSA-proteolytic enzyme.  
 DE DE  
 XX Cleavage site; prostate specific antigen; PSA; proteolytic enzyme;  
 KW produg; prostatic glandular cell; sesquiterpene-gamma-lactone;  
 XX cell proliferation; prostate cancer; breast cancer.  
 OS Synthetic.  
 XX WO9852966-A1.  
 PN PD  
 XX 26-NOV-1998.  
 PD PD  
 XX 19-MAY-1998; 98WO-US10285.  
 PF PF  
 XX 30-MAR-1998; 98US-0080046.  
 PR PR  
 XX 19-MAY-1997; 97US-0047070.  
 PR PA  
 XX (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
 XX Christensen SB, Denmeade SR, Issacs JT, Lilja H;  
 PI WPI; 1999-131688/11.  
 XX Peptides used in compositions to treat prostate-specific  
 PT antigen-producing cell proliferation - have amino-acid sequence with  
 PT cleavage site specific for PSA-proteolytic enzyme  
 XX Claim 8; Page 44; 58pp; English.  
 PS The present sequence represents a peptide comprising an amino-acid  
 CC sequence with a cleavage site specific for an enzyme with the  
 CC proteolytic activity of prostate-specific antigen (PSA). The peptide  
 CC can be used in compositions for the treatment of PSA-producing cell  
 CC proliferative disorders, both benign and malignant, such as prostate and  
 CC breast cancers, to detect PSA-producing tissue, to select PSA-  
 CC activatable produgs substantially specific for target tissue producing  
 CC PSA-producing cells to determine activity of PSA in samples and to image  
 CC PSA-producing tissue. The peptide can be linked to therapeutic drugs,  
 CC such as thapsigargin, estafiatin, grossheimin, inuchineolide, arglavin  
 CC and derivatives such as thapsigargin, to form produgs thereby

CC substantially inhibiting non-specific toxicity by activating drug  
CC locally at PSA-producing when enzymatic cleavage takes place.

XX  
SQ Sequence 7 AA;

Query Match 26.7%; Score 4; DB 20; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HSKK 4  
    ||||  
Db 2 hskk 5

Search completed: September 15, 2001, 12:44:28  
Job time: 314 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 15, 2001, 12:53:46 : Search time 11.49 Seconds  
(without alignments)  
43.009 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_163  
Perfect score: 24  
Sequence: 1 HSKSEKLRERIRKVCCEQLRTLLP 24

Scoring table:  
Gapop 60.0, Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 114941

Minimum DB seq length: 0  
Maximum DB seq length: 24

Post-processing: Listing first 45 summaries

Database : Issued Patents-AA:\*  
1: /cgnl\_7/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgnl\_7/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgnl\_7/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgnl\_7/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgnl\_7/ptodata/1/1aa/PCtus.COMB.pep:\*  
6: /cgnl\_7/ptodata/1/1aa/Backfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	25.0	15	4	US-08-986-837-8
2	5	20.8	6	1	US-07-944-143C-17
3	5	20.8	6	5	PCT-US93-08214-17
4	5	20.8	6	5	PCT-US93-12679-5
5	5	20.8	7	1	US-08-105-416-36
6	5	20.8	7	2	US-08-473-656A-36
7	5	20.8	10	1	US-08-222-851-14
8	5	20.8	10	1	US-08-440-504A-9
9	5	20.8	10	1	US-08-440-504A-14
10	5	20.8	10	5	PCT-US94-12985-13
11	5	20.8	14	4	US-08-433-613-9
12	5	20.8	14	4	US-08-433-613-48
13	5	20.8	16	1	US-07-944-143C-16
14	5	20.8	16	5	PCT-US93-08214-16
15	5	20.8	17	1	US-07-976-872B-1
16	5	20.8	17	1	US-07-976-872B-2
17	5	20.8	17	1	US-07-976-872B-3
18	5	20.8	17	1	US-07-976-872B-4
19	5	20.8	17	1	US-08-105-416-12
20	5	20.8	17	1	US-08-105-416-13
21	5	20.8	17	1	US-08-105-416-18
22	5	20.8	17	1	US-08-105-416-19
23	5	20.8	17	1	US-08-105-416-20
24	5	20.8	17	1	US-08-105-416-27
25	5	20.8	17	1	US-08-105-416-29
26	5	20.8	17	1	US-08-105-416-30
27	5	20.8	17	1	US-08-105-416-37

28	5	20.8	17	1	US-08-105-416-38	Sequence 38, Appl
29	5	20.8	17	2	US-08-473-656A-12	Sequence 12, Appl
30	5	20.8	17	2	US-08-473-656A-13	Sequence 13, Appl
31	5	20.8	17	2	US-08-473-656A-18	Sequence 18, Appl
32	5	20.8	17	2	US-08-473-656A-19	Sequence 19, Appl
33	5	20.8	17	2	US-08-473-656A-20	Sequence 20, Appl
34	5	20.8	17	2	US-08-473-656A-27	Sequence 27, Appl
35	5	20.8	17	2	US-08-473-656A-29	Sequence 29, Appl
36	5	20.8	17	2	US-08-473-656A-30	Sequence 30, Appl
37	5	20.8	17	2	US-08-473-656A-37	Sequence 37, Appl
38	5	20.8	17	2	US-08-473-656A-38	Sequence 38, Appl
39	5	20.8	17	3	US-08-483-931B-12	Sequence 12, Appl
40	5	20.8	17	3	US-08-483-931B-13	Sequence 13, Appl
41	5	20.8	17	3	US-08-483-931B-18	Sequence 18, Appl
42	5	20.8	17	3	US-08-483-931B-19	Sequence 19, Appl
43	5	20.8	17	3	US-08-483-931B-20	Sequence 20, Appl
44	5	20.8	17	3	US-08-483-931B-27	Sequence 27, Appl
45	5	20.8	17	3	US-08-483-931B-29	Sequence 29, Appl

## ALIGNMENTS

RESULT 1  
US-08-986-837-8  
; Sequence 8, Application US/08986837  
; Patent No. 6221676  
; GENERAL INFORMATION:  
; APPLICANT: Lam, Bing K.  
; APPLICANT: Penrose, John F.  
; APPLICANT: Frank, Austen K.  
; TITLE OF INVENTION: DNA ENCODING HUMAN LEUKOTRIENE C4 SYNTHASE,  
; FILE REFERENCE: POLYPEPTIDES, AND USES THEREFOR  
; CURRENT APPLICATION NUMBER: US/08/986,837  
; EARLIER FILING DATE: 1997-12-18  
; EARLIER APPLICATION NUMBER: 08/246,991  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-986-837-8

Query Match 25.0%; Score 6; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLLP 24  
DB 8 LRTLLP 13

RESULT 2  
US-07-944-143C-17  
; Sequence 17, Application US/07944143C  
; Patent No. 5719064  
; GENERAL INFORMATION:  
; APPLICANT: Scofield, R. Hal  
; APPLICANT: Hartley, John B.  
; TITLE OF INVENTION: Peptide Diagnostics and Therapeutics for  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patricia L. Pabst  
; STREET: 2800 One Atlantic Center  
; CITY: Atlanta  
; STATE: GA  
; COUNTRY: USA

ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/944,143C  
FILING DATE:  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)873-8794  
TELEFAX: (404)873-8795  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
ORGANISM: Salmonella typhimurium  
US-07-944-143C-17

Query Match 20.8%; Score 5; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
DB 2 LRTLL 6

RESULT 3  
PCT-US93-08214-17  
Sequence 17, Application PC/TUS9308214  
GENERAL INFORMATION:  
APPLICANT: Oklahoma Medical Research Foundation  
TITLE OF INVENTION: Peptide Diagnostics and Therapeutics  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: United States  
ZIP: 30309-4530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08214  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 815-6508  
TELEFAX: (404) 815-6555  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
for Spondyloarthropat

LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
ORGANISM: Salmonella typhimurium  
PCT-US93-08214-17

Query Match 20.8%; Score 5; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
DB 2 LRTLL 6

RESULT 4  
PCT-US93-12679-5  
Sequence 5, Application PC/TUS9312679  
GENERAL INFORMATION:  
APPLICANT: Rath, Mathias  
TITLE OF INVENTION: SYNTHETIC OLIGOPEPTIDES ANALOGOUS TO  
TITLE OF INVENTION: PROTEIN SIGNAL SEQUENCES -METHODS OF USE  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SHELDON & MAX  
STREET: 401 Florence Street  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/12679  
FILING DATE: 30-DEC-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cranfill, Raymond B  
REGISTRATION NUMBER: 32,845  
REFERENCE/DOCKET NUMBER: RATH-10016PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-322-5333  
TELEFAX: 415-322-5499  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US93-12679-5

Query Match 20.8%; Score 5; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKER 6  
DB 2 SSKER 6

RESULT 5



US-08-105-416-36  
: Sequence 36, Application US/08105416  
: Patent No. 5639958  
: GENERAL INFORMATION:  
: APPLICANT: Olsson, Lennart  
: APPLICANT: Goodenow, Robert S  
: APPLICANT: Goldstein, Avram  
: TITLE OF INVENTION: Class I MHC Modulation of Surface  
: TITLE OF INVENTION: Receptor Activity  
: NUMBER OF SEQUENCES: 40  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
: STREET: Four Embarcadero Center, Suite 3400  
: CITY: San Francisco  
: STATE: CA  
: COUNTRY: US  
: ZIP: 94111  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/105,416  
: FILING DATE: 12-AUG-1993  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Rowland, Berttram I.  
: REGISTRATION NUMBER: 20015  
: REFERENCE/DOCKET NUMBER: A55115-4/BIR  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 415-781-1989  
: TELEFAX: 415-398-3249  
: INFORMATION FOR SEQ ID NO: 36:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 7 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: peptide  
: US-08-105-416-36

Query Match 20.8%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
|||||  
DB 1 LRTLL 5

RESULT 6  
US-08-473-656A-36  
: Sequence 36, Application US/08473656A  
: Patent No. 5853999  
: GENERAL INFORMATION:  
: APPLICANT: Olsson, Lennart  
: APPLICANT: Goodenow, Robert S  
: APPLICANT: Goldstein, Avram  
: TITLE OF INVENTION: Class I MHC Modulation of Surface  
: TITLE OF INVENTION: Receptor Activity  
: NUMBER OF SEQUENCES: 40  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
: STREET: Four Embarcadero Center, Suite 3400  
: CITY: San Francisco  
: STATE: CA  
: COUNTRY: US  
: ZIP: 94111  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/473,656A  
: FILING DATE:  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/105,416  
: FILING DATE: 12-AUG-1993  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Rowland, Berttram I.  
: REGISTRATION NUMBER: 20015  
: REFERENCE/DOCKET NUMBER: A55115-4/BIR  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 415-781-1989  
: TELEFAX: 415-398-3249  
: INFORMATION FOR SEQ ID NO: 36:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 7 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: peptide  
: US-08-473-656A-36

Query Match 20.8%; Score 5; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
|||||  
DB 1 LRTLL 5

RESULT 7  
US-08-222-851-14  
: Sequence 14, Application US/08222851  
: Patent No. 5723128  
: GENERAL INFORMATION:  
: APPLICANT: CLAYBERGER, CAROL A.  
: APPLICANT: KRENSKY, ALAN M.  
: APPLICANT: PARHAM, PETER  
: TITLE OF INVENTION: CYTOTOXIC T-CELL LYMPHOCYTE ("CTL")  
: TITLE OF INVENTION: ACTIVITY REGULATION BY CLASS I MHC PEPTIDES  
: NUMBER OF SEQUENCES: 43  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: MORRISON & FOERSTER  
: STREET: 2000 PENNSYLVANIA AVENUE, NW, STE 5500  
: CITY: WASHINGTON  
: STATE: DC  
: COUNTRY: USA  
: ZIP: 20006-1812  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/222,851  
: FILING DATE: 05-APR-1994  
: CLASSIFICATION: 424  
: ATTORNEY/AGENT INFORMATION:  
: NAME: MILLMAN, ROBERT A.  
: REGISTRATION NUMBER: 36,217  
: REFERENCE/DOCKET NUMBER: 28600-20200.22  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (202) 887-1500  
: TELEFAX: (202) 494-0792  
: TELEX: 90-4030 MRSNFOERSMHS  
: INFORMATION FOR SEQ ID NO: 14:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 10 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-222-851-14

Query Match 20.8%; Score 5; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLL 23  
|||||  
DB 4 LRTLL 8

RESULT 8  
US-08-440-504A-9  
Sequence 9, Application US/08440504A  
Patent No. 5753625

GENERAL INFORMATION:  
APPLICANT: BueLOW, Roland  
TITLE OF INVENTION: Treatment for Inhibiting the Progression  
of Autoimmune Disease  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESS: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA

COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440.504A  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-60130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 949-8711  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-440-504A-9

Query Match 20.8%; Score 5; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLL 23  
|||||  
DB 4 LRTLL 8

RESULT 9  
US-08-440-504A-14  
Sequence 14, Application US/08440504A  
Patent No. 5753625

GENERAL INFORMATION:  
APPLICANT: BueLOW, Roland  
TITLE OF INVENTION: Treatment for Inhibiting the Progression

TITLE OF INVENTION: of Autoimmune Disease  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESS: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440.504A  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-60130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 949-8711  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-440-504A-14

Query Match 20.8%; Score 5; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLL 23  
|||||  
DB 4 LRTLL 8

RESULT 10  
PCT-US94-12985-13  
Sequence 13, Application PC/TUS9412985

GENERAL INFORMATION:  
APPLICANT: The Board of Trustees for the Leland Stanford Junior  
UNIVERSITY  
TITLE OF INVENTION: SURFACE MEMBRANE PROTEINS AND THEIR  
EFFECT ON IMMUNE RESPONSE  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESS: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/12985  
FILING DATE: 10-NOV-1994  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/150,493  
FILING DATE: 10-NOV-1993  
ATTORNEY/AGENT INFORMATION:

Query Match 20.8%; Score 5; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLL 23  
|||||  
DB 4 LRTLL 8

RESULT 10  
PCT-US94-12985-13  
Sequence 13, Application PC/TUS9412985

GENERAL INFORMATION:  
APPLICANT: The Board of Trustees for the Leland Stanford Junior  
UNIVERSITY  
TITLE OF INVENTION: SURFACE MEMBRANE PROTEINS AND THEIR  
EFFECT ON IMMUNE RESPONSE  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESS: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/12985  
FILING DATE: 10-NOV-1994  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/150,493  
FILING DATE: 10-NOV-1993  
ATTORNEY/AGENT INFORMATION:

NAME: Rowland, Bertram I  
REGISTRATION NUMBER: 20,015  
REFERENCE/DOCKET NUMBER: FP-58976-PC/BIR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US94-112985-13

Query Match 20.8%; Score 5; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
|||||  
DB 4 LRTLL 8

RESULT 11  
US-08-433-613-9  
Sequence 9, Application US/08433613A  
Patent No. 6162434  
GENERAL INFORMATION:  
APPLICANT: Buelow, Roland  
TITLE OF INVENTION: Cytomodulating Peptide for Inhibiting Lymphocyte  
FILE REFERENCE: A61008/RET/TAL  
CURRENT APPLICATION NUMBER: US/08/433,613A  
CURRENT FILING DATE: 1995-05-03  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-08-433-613-9

Query Match 20.8%; Score 5; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
|||||  
DB 8 LRTLL 12

RESULT 12  
US-08-433-613-48  
Sequence 48, Application US/08433613A  
Patent No. 6162434  
GENERAL INFORMATION:  
APPLICANT: Buelow, Roland  
TITLE OF INVENTION: Cytomodulating Peptide for Inhibiting Lymphocyte  
FILE REFERENCE: A61008/RET/TAL  
CURRENT APPLICATION NUMBER: US/08/433,613A  
CURRENT FILING DATE: 1995-05-03  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 48  
LENGTH: 14  
TYPE: PRT

ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-08-433-613-48

Query Match 20.8%; Score 5; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
|||||  
DB 8 LRTLL 12

RESULT 13  
US-07-944-143C-16  
Sequence 16, Application US/07944143C  
Patent No. 5719064  
GENERAL INFORMATION:  
APPLICANT: Scofield, R. Hal  
APPLICANT: Harley, John B.  
TITLE OF INVENTION: Peptide Diagnostics and Therapeutics for  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/944,143C  
FILING DATE:  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)873-8794  
TELEFAX: (404)873-8795  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
US-07-944-143C-16

Query Match 20.8%; Score 5; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
|||||  
DB 11 LRTLL 15

RESULT 14  
PCT-US93-08214-16

Sequence 16, Application PC/TUS9308214  
GENERAL INFORMATION:  
APPLICANT: Oklahoma Medical Research Foundation  
TITLE OF INVENTION: Peptide Diagnostics and Therapeutics for Spondyloarthropat  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: United States  
ZIP: 30309-4530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08214  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRFL38  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 815-6508  
TELEFAX: (404) 815-6555  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
PCT-US93-08214-16

Query Match 20.8%; Score 5; DB 5; Length 16;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 19 LRTLL 23  
|||||  
Db 11 LRTLL 15

RESULT 15  
US-07-976-872B-1  
Sequence 1, Application US/07976872B  
Patent No. 5516642  
GENERAL INFORMATION:  
APPLICANT: Mapelli, Claudio  
APPLICANT: Meyers, Chester A.  
TITLE OF INVENTION: POLYPEPTIDES DERIVED FROM MAJOR  
TITLE OF INVENTION: HISTOCOMPATIBILITY COMPLEX CLASS I ANTIGEN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burton Rodney  
STREET: P.O. Box 4000  
CITY: Princeton  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 08543-4000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/976,872B  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gaul, Timothy J.  
REGISTRATION NUMBER: 33,111  
REFERENCE/DOCKET NUMBER: KX41  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (609) 252-5901  
TELEFAX: (609) 252-4526  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-976-872B-1

Query Match 20.8%; Score 5; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 19 LRTLL 23  
|||||  
Db 10 LRTLL 14

Search completed: September 15, 2001, 12:55:31  
Job time: 105 sec

Sun Sep 16 18:13:07 2001

us-09-389-000-2\_copy\_140\_163.tif

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:53:01 ; Search time 17.48 Seconds  
(without alignments)  
83.237 Million cell updates/sec

Title:	US-09-389-000-2_COPY_140_163
Perfect score:	24
Sequence:	1 HSKSEKLRRIKVCCEQLRTLLP 24

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

```
Searched: 412676 seqs, 60623988 residues
Word size : 0
```

Total number of hits satisfying chosen parameters: 172165

```
Minimum DB seq length: 0
Maximum DB seq length: 24
```

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_0601:\*

1:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1980.DAT *
2:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1981.DAT *
3:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1982.DAT *
4:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1983.DAT *
5:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1984.DAT *
6:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1985.DAT *
7:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1986.DAT *
8:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1987.DAT *
9:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1988.DAT *
10:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1989.DAT *
11:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1990.DAT *
12:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1991.DAT *
13:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1992.DAT *
14:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1993.DAT *
15:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1994.DAT *
16:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1995.DAT *
17:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1996.DAT *
18:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1997.DAT *
19:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1998.DAT *
20:	/cgnl_9/gcgdata/geneseq/geneseqp/AA1999.DAT *
21:	/cgnl_9/gcgdata/geneseq/geneseqp/AA2000.DAT *
22:	/cgnl_9/gcgdata/geneseq/geneseqp/AA2001.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14	58.3	15	21	AAV79270	PHLIX peptide used
2	5	20.8	6	4	AA30395	Sequence of Aas 26
3	5	20.8	6	15	AA50267	Peptide correspond
4	5	20.8	6	15	AA59948	Peptide signal seq
5	5	20.8	9	15	AA61019	Polyspharin-like pol
6	5	20.8	9	16	AA84869	H2kb 75-83 immuno
7	5	20.8	10	14	AA41212	Peptide fragment c
8	5	20.8	10	16	AA89543	HLA-B2705.75-84.
9	5	20.8	10	16	AA83075	HLA-B2702 CTL mod
10	5	20.8	10	17	AAW07524	T-cell modulating
11	5	20.8	10	17	AAW07517	T-cell modulating

12	10	19	AAM33785
13	5	20.8	Pepide B2705_75-8
13	5	20.8	Immunosuppressive
14	5	20.8	Immunosuppressive
14	5	20.8	Glucose transport
15	5	20.8	Guanine thymine b1
16	5	20.8	Human Class I HLA-
17	5	20.8	Human Class I HLA-
18	5	20.8	Oestrogen receptor
19	5	20.8	HLA B27 hypervaria
20	5	20.8	Glucose transport
21	5	20.8	Human MHC 1 and HL
22	5	20.8	Human HLA-B27-(62
23	5	20.8	Human [Phe74]-HLA-
24	5	20.8	Human MHC 1 alpha
25	5	20.8	Human MHC 1 alpha
26	5	20.8	Human MHC 1 alpha
27	5	20.8	Human MHC 1 alpha
28	5	20.8	Human MHC 1 alpha
29	5	20.8	Human MHC 1 alpha
30	5	20.8	MHC peptide repeat
31	5	20.8	MHC peptide repeat
32	5	20.8	MHC peptide repeat
33	5	20.8	Pepide membrane b
34	5	20.8	p1 swap 2 muclein c
35	5	20.8	UL snRNP 70K proteo
36	5	20.8	Human neuroendocr
37	5	20.8	Human MHC alpha-1
38	5	20.8	Human MHC Class I-
39	5	20.8	Human MHC 1 alpha
40	5	20.8	Human MHC 1 alpha
41	5	20.8	Human MHC 1 alpha
42	5	20.8	Human MHC 1 alpha
43	5	20.8	Human MHC 1 alpha
44	5	20.8	GPV Landeram Leu-ri
45	5	20.8	peptide targeted

## ALIGNMENTS

## RESULT 1

ID: AAY79270 standard; Peptide; 15 AA.

AC AAY79270;

DT 03-JUL-2000 (first entry)

DE PHELIX peptide used to raise antibody.

KM PHELI-X, human; testis-specific; transcription factor;  
KM prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
KW therapy; diagnosis; vaccine; antibody.

OS Homo sapiens.

PN WO200012709-A2

09-MAR-2000  
PD

PF 31-AUG-1999; 99WO-US20137.

PR 31-AUG-1998; 98US-0098610.

PA (UROG-) UROGENESYS INC.

PA (HUBE/) HUBERT R S.

XX

DR WPI; 2000-237872/20

XX

PT Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -

XX Example 5; Page 36; 62pp; English.

CC The present sequence is that of a peptide derived from human  
 CC PHELIIX (see AAY79269), a novel transcription factor that is  
 CC normally expressed only in testis tissue, but which is up-regulated  
 CC in prostate and some other cancers. The peptide was conjugated to  
 CC keyhole limpet haemocyanin and used to raise polyclonal antiserum  
 CC in rabbit. The antiserum demonstrated specificity for PHELIIX and  
 CC may therefore be useful for assessing the expression of PHELIIX in  
 CC patient samples.

SO Sequence 15 AA;

Query Match 58.3%; Score 14; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 7e-08;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HSSKEKLRERIKY 14  
 1 HSSKEKLRERIKY 14

Db

RESULT 2

AAP30295  
 ID AAP30295 standard; Protein: 6 AA.

XX AAP30295;

XX 20-APR-1992 (first entry)

DE Sequence of AAs 26-31 of the E. coli heat stable toxin  
 DE which determine an H-peptide.

XX Synthetic vaccine; antigen; allergen; immunological response;  
 KW antibody.

XX Escherichia coli.

OS Ep93851-A.

PN 16-NOV-1983.

XX 11-MAR-1983; 83EP-0102392.

XX 15-MAR-1982; 82US-0358150.

PR 28-JAN-1983; 83US-0461802.

PR 12-JUN-1981; 81US-0272855.

PR 09-JAN-1981; 81US-0233558.

PR 16-DEC-1986; 86US-0942562.

XX (NYBL-) NEW YORK BLOOD CENT.

PA Hopp TP;

XX WPI: 1983-822049/47.

XX Synthetic vaccine - contains peptide residue coupled to higher  
 PT alkyl or alkenyl Gps. and with 6 amino acids in residue

XX Claim 50; Page 46; 54pp; English.

XX The inventory claim a synthetic vaccine which comprises a peptide  
 CC residue coupled to an alkyl or alkenyl gp. having at least 12C, or  
 CC other lipophilic substance. The residue contains a sequence of 6 AAs  
 CC corresp. to the SO of such AAs in a protein antigen or allergen where  
 CC the greatest local average hydrophobicity is found. Pref. the AAs in  
 CC the peptide do not exceed 50 residues, and they esp. contain 12-18  
 CC residues. The alkyl or alkenyl gp. pref. contains 12-24C and it is

CC pref. coupled to the terminal amino gp. of the residue opt. via a CO  
 CC gp. If a lipophilic substance is used, it is pref. palmitic, stearic,  
 CC benenic, oleic or mycolic acid.

SO Sequence 6 AA;

Query Match 20.8%; Score 5; DB 4; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSKK 6  
 2 SSKK 6

Db

RESULT 3

AAR50267  
 ID AAR50267 standard; peptide: 6 AA.

XX AAR50267;

XX 13-OCT-1994 (first entry)

DE Peptide corresponding to a shared sequence of an HLA molecule.

XX Peptide; HLA; human leucocyte antigen; diagnosis; treatment;

KW autoimmune disease; uveitis; spondylitis; psoriasis;

KW inflammatory bowel disease; enteric bacteria;

XX Salmoneilla typhimurium.

OS Salmoneilla typhimurium.

PN WO9405303-A.

XX 17-MAR-1994.

XX 31-AUG-1993; 93WO-US08214.

PR 31-AUG-1992; 92US-0944143.

XX (OKLA ) UNIV OKLAHOMA STATE.

PA Harley JB, Scofield RH;

PI WPI: 1994-100843/12.

XX Peptide corresponding to shared sequences of HLA molecules and  
 PT enteric bacteria - used for the diagnosis and treatment of  
 PT auto-immune disorders, partic. spondylarthropathies.

XX Disclosure; Page 15; 58pp; English.

XX The peptide is a fragment of the HLA B27 hypervariable region  
 CC described in AAR50266. The peptide can be used for the diagnosis and  
 CC treatment of autoimmune disorders, e.g. spondyloarthropies including  
 CC uveitis and spondylitis associated with inflammatory bowel disease  
 CC or psoriasis.

XX Sequence 6 AA;

Query Match 20.8%; Score 5; DB 15; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLL 23  
 2 LRTLL 6

Db

RESULT 4

AAR59948



ID AAR59948 standard; peptide: 6 AA.  
 XX AAR59948;  
 AC  
 XX 14-FEB-1995 (first entry)  
 DT  
 XX Peptide signal sequence for treating E. coli infections.  
 DE  
 XX Therapeutic; metabolic interactions; PSS; analogues.  
 KW  
 XX Synthetic.  
 OS  
 XX W09416328-A.  
 PN  
 XX 21-JUL-1994.  
 PD  
 XX 30-DEC-1993; 93WO-US12679.  
 PF  
 XX 30-DEC-1992; 92US-0997727.  
 PR  
 XX (RATH/) RATH M.  
 PA  
 XX Rath M.  
 PI  
 XX WPI; 1994-249399/30.  
 DR  
 XX Identifying peptide signal sequences in a protein - and use of  
 PT their synthetic analogues for treating or preventing, e.g.  
 PT cardiovascular and auto-immune disease, infections and cancer.  
 XX  
 PS Claim 18; Page 10; 28pp; English.  
 XX  
 CC The sequence is that of a peptide signal sequence which can be used  
 CC to treat E. coli infections.  
 CC See also AAR59944-83.  
 CC  
 SO Sequence 6 AA;

OY 2 SSKEX 6  
 DB 2 sskex 6

Query Match 20.8%; Score 5; DB 15; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5  
 AAR61019  
 ID AAR61019 standard; peptide: 9 AA.  
 XX  
 AC AAR61019;  
 AC  
 XX 23-APR-1995 (first entry)  
 DT  
 XX Dynorphin-like polypeptide.  
 DE  
 XX Dynorphin; opioid analgesic; stable; stability; intravenous.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1 /note= "MetYr"  
 FT Misc-difference 8 /note= "D-Glu"  
 FT Modified-site 9 /note= "Arg-NH2"  
 FT  
 XX EP614913-A.  
 PN  
 XX 14-SEP-1994.  
 PD

XX 08-NOV-1985; 94EP-0107769.  
 PE  
 XX 09-NOV-1984; 84JP-0236076.  
 PR  
 XX (EISA ) EISAI CO LTD.  
 PA  
 XX Arakawa Y, Araki S, Ikeda M, Kaneko T, Nakazawa T;  
 PI Tachibana S, Tsuchiya Y, Yamatsu K, Yoshino H;  
 XX WPI; 1994-281157/35.  
 DR  
 XX  
 PT New dynorphin polypeptide derivs. for use as analgesics - contg.  
 PT a D-amino acid or an N-alkyl deriv. of a D- or L-amino acid for  
 PT in vivo stability.  
 XX  
 PS Example 26; Page 27; 47pp; English.  
 XX  
 CC The patent discloses dynorphin-like polypeptides which are shorter  
 CC than dynorphin (7 to 9 amino acids) and which contain either a D-  
 CC amino acid and/or an N-alkyl deriv. of an L- or D-amino acid to  
 CC provide in-vivo stability upon intravenous administration. The  
 CC peptides have a much greater analgesic effect than dynorphin  
 CC due to their greater stability in the blood (e.g. IC50 values can be  
 CC as low as 0.04 nM when tested by the rabbit vas deferens method; cf.  
 CC 17.4 nM for dynorphin).  
 CC The present peptide is a specific example of the peptides.  
 CC  
 SQ Sequence 9 AA;

OY 7 LRRER 11  
 DB 5 lrrer 9

Query Match 20.8%; Score 5; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6  
 AAR84869  
 ID AAR84869 standard; peptide: 9 AA.  
 XX  
 AC AAR84869;  
 AC  
 XX 30-MAY-1996 (first entry)  
 DT  
 XX H2KB 75-83 immunogenic peptide.  
 DE  
 XX Antigen; epitope; cell mediated; immune specific; cancer;  
 KW infection; infestation; mucin-1; MUC-1; tumour; H2KB 75-83;  
 KW immunogenic peptide.  
 KW  
 XX Synthetic.  
 OS  
 XX W09527505-A1.  
 PN  
 XX 19-OCT-1995.  
 PD  
 XX 12-APR-1995; 95WO-US04540.  
 PF  
 XX 12-APR-1994; 94US-0229606.  
 PR  
 XX (BIOM-) BIOMIRA INC.  
 PA  
 XX Ding L, Koganty RR, Longenecker BM, Reddish MA;  
 PI WPI; 1995-373528/48.  
 DR  
 XX New cell-mediated immune-specific immunogenic compsns. - used in  
 PT prophylaxis and treatment of cancer, microbial infections, viral  
 PT infections and parasitic infestations

Example 1: Page 62; 141pp; English.

A claimed cell-mediated immune (CMI)-specific immunogenic compsn. comprises a conjugate of a primary antigen bearing a primary epitope, with an immunomodulatory peptide (IP), i.e. the present peptide. The IP comprises an allotopeptide moiety of at least 5 amino acids, whose sequence corresponds to a polymorphic region of a MHC encoded polymorphic Class I or II antigen. The compsn. can be used to elicit a CMI-specific response which is prophylactic, or therapeutic for, e.g. microbial and viral infections, parasitic infections and cancer, partic. MUC-1 expressing tumor cells when the present peptide is the IP, and a MUC-1 epitope is the primary epitope.

Sequence 9 AA:

Query Match 20.8%; Score 5; DB 16; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

19 LRTL 23  
| | | | |  
4 LRTL 8

RESULT 7  
AAR41212  
ID AAR41212 standard; peptide; 10 AA.  
AAR41212;  
15-MAR-1994 (first entry)  
Peptide fragment of Class I HLA peptide.  
Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;  
parasitic disease; cytotoxic T lymphocyte; modulation.  
Synthetic.  
W09317699-A.  
16-SEP-1993.  
25-FEB-1993; 93WO-US01758.  
02-MAR-1992; 92US-0844716.  
(STRD ) UNIV LELAND STANFORD JUNIOR.  
Clayberger CA, Krensky AM;  
WPI; 1993-303134/38.  
New peptide(s) based on Class I HLA antigen domains - used for modulating cytotoxic T-lymphocyte activity towards targets  
Claim 11; Page 54; 61pp; English.

The peptide is used to modulate cytotoxic T-lymphocyte (CTL) activity, either by inhibition or stimulation. It can be used for inhibiting CTL toxicity in transplantations, for inducing CTL activity in parasitic diseases and neoplasia and in studies on viral infection. The peptide can also be used for identifying CTLs which bind to it and removing subsets of CTLs from a T-cell composition. This peptide sequence is more commonly found within larger peptide compounds of not more than 30 amino acids in length.

Sequence 10 AA:

Query Match	20.8%	Score 5;	DB 14;	Length 10;
Best Local Similarity	100.0%;	Pred. No. 62;		
Matches 5;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	19	lrrtl	23	.
Db	4	lrrtl	8	

RESULT	8
AAR95423	
ID	AAR95423 standard; peptide; 10 AA
XX	
AC	AAR95423;
XX	
DT	12-NOV-1996 (first entry)

HLA-B2705.75-84.

HLA; p74; alpha1-helix; human-leucocyte-associated antigen; inhibitor; T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC; B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation; cytolysis; antigen presenting cell.

Synthetic.

W09513288-A1

18-MAY-1995.

10-NOV-1994; 94WO-US12985.

10-NOV-1993; 93US-0150493.

(STRD ) UNIV LELAND STANFORD JUNIOR.

Clayberger C, Krensky AM;

WPI; 1995-194027/25.

Compsns. comprising lymphoid surface membrane proteins - which may inhibit cytolytic activity and differentiation of CTLs.

Example; Page 11; 29pp; English.

AA95411 and AA95415-95431 represent palindromes and fragments of human-Leucocyte-associated antigens. This sequence represents the HLA-B\*2705.75-84. These sequences can be used to isolate the protein p74 from a T-cell lysate. p74 is a T-cell surface membrane protein associated with T-cell activation in mammalian T-cells, and is also immunologically cross reactive with the heat shock protein Hsc70. p74 is found in a limited number of cell types, but is particularly expressed on B and T cells. p74 can be isolated by lysis of a suitable cell with an amphipathic detergent, and then passed through an affinity column containing a covalently bound HLA-B\*2702 palindromic peptide. Compositions comprising the extracellular fragment of p74 combined with HLA-B\*2702.60-84 (see AA95416), induces calcium influx, and inhibits cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate compounds can be screened for their effect on the cyolytic activity of T-cells, by combining them with the extracellular portion of p74 and determining the amount of binding between the candidate compound and p74. Modulation of CTL activity can be inhibited in a cellular composition containing T-cells and antigen presenting cells (APCs), by adding to the mix the extracellular portion of p74, in an amount sufficient to compete with p74 for the binding of the p74 ligand.

Query Match	20.8%;	Score 5;	DB 16;	Length 10;
Best Local Similarity	100.0%;	Pred. No. 62;		
Matches	5;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

OY 19 LRTLL 23  
 |||||  
 DB 4 lrtll 8

## RESULT 9

AAW83075  
 ID AAR83075 standard; peptide; 10 AA.

XX AAR83075;

XX 16-MAY-1996 (first entry)

DE HLA-B2702 CTL modulating peptide (B2702.75-84).

XX Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW immunosuppressant; graft versus host disorder; transplantation; therapy;  
 KW class I MHC; HLA-B2702.

XX Synthetic.

XX WO9526979-A1.

XX 12-OCT-1995.

XX 05-APR-1995; 95WO-US04349.

XX 05-APR-1994; 94US-0222851.

XX (STFRD ) UNIV LELAND STANFORD JUNIOR.

XX Clayberger C, Krensky AM, Parham P;

XX WPI; 1995-358582/46.

PT Extension of acceptance period of transplants from MHC unmatched  
 PT donor hosts - using Class I B75-84 MHC antigen of the recipient  
 PT host

XX Example 14; Page 34; 80pp; English.

XX AAR83061-R83085, AAR83090-R83096 and AAR92907-R92913 represent fragments  
 CC of class I major histocompatibility complex (MHC) antigens. This  
 CC sequence corresponds to residues 75-84 of the alpha-1 domain of the  
 CC class I MHC HLA-B2702. These sequences can be used to extend the period  
 CC of acceptance by a recipient of a transplant from an MHC unmatched  
 CC donor. The peptides are administered to a patient in conjunction with a  
 CC subtherapeutic amount of an immunosuppressant. This is administered to  
 CC the patient for a limited period of time (compared to the lifetime  
 CC administration for current treatments). The peptides particularly  
 CC modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 CC of the patient.

XX Sequence 10 AA;

Query Match 20.8%; Score 5; DB 16; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLL 23  
 |||||  
 DB 4 lrtll 8

## RESULT 10

AAW07524  
 ID AAW07524 standard; peptide; 10 AA.

XX AAW07524;

XX 04-AUG-1997 (first entry)

XX

DE T-cell modulating peptide DK.

XX T-cell modulator; autoimmune disease; tissue destruction; alpha-domain;  
 KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;  
 KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;  
 KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;  
 KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;  
 KW autologous target cell; cytokine release; T cell activation; therapy.

XX Synthetic.

XX WO9635443-A1.

XX 14-NOV-1996.

XX 05-APR-1996; 96WO-US04710.

XX 12-MAY-1995; 95US-0440504.

XX (SANG-) SANGSTAT MEDICAL CORP.

XX Buelow R;

XX WPI; 1996-518410/51.

PT Treatment of auto-immune disease by admin. of peptide(s) corresp. to  
 PT major histocompatibility complex antigens - esp. for delaying onset  
 PT of clinical symptoms of insulin dependent diabetes by modulating T  
 PT cell mediated attack on target cells

XX Example 1; Page 12; 24pp; English.

XX AAW07521-W07524, and AAW07527 represent T-cell modulating peptides that  
 CC can be used in the method of the invention. These sequences are based on  
 CC a portion of the generic peptide corresponding to residues 70-91 of the  
 CC alpha-domain of the major histocompatibility complex (MHC) class I  
 CC antigen (see AAW07510). The method is for affecting the course of an  
 CC autoimmune disease involving T-cell mediated destruction of tissue in  
 CC mammals. These peptides are used especially to treat insulin-dependent  
 CC diabetes mellitus, preferably being administered during the pre-clinical  
 CC stage to delay onset of the disease. Other diseases that can be treated  
 CC are multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus  
 CC vulgaris, Sjogren's disease, thyroid disease, Hashimoto's thyroiditis,  
 CC myasthenia gravis, etc. The peptides modulate T-cell mediated attack on  
 CC autologous target cells, and may also reduce inflammation, swelling, and  
 CC release of cytokines, perforins, granzymes etc. associated with T cell  
 CC activation.

XX Sequence 10 AA;

Query Match 20.8%; Score 5; DB 17; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLL 23  
 |||||  
 DB 4 lrtll 8

## RESULT 11

AAW07517  
 ID AAW07517 standard; peptide; 10 AA.

XX AAW07517;

XX 04-AUG-1997 (first entry)

XX T-cell modulating peptide #6.

XX T-cell modulator; autoimmune disease; tissue destruction; alpha-domain;  
 KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;  
 KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;

KM Rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;  
 KM thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;  
 KM autologous target cell; cytokine release; T cell activation; therapy.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9635443-A1.  
 XX  
 PD 14-NOV-1996.  
 XX  
 PE 05-APR-1996; 96WO-US04710.  
 XX  
 PR 12-MAY-1995; 95US-0440504.  
 XX  
 PA (SANGC-) SANGSTAT MEDICAL CORP.  
 XX  
 PI Buelow R;  
 XX  
 DR WPI: 1996-518410/51.  
 XX  
 PT Treatment of auto-immune disease by admin. of peptide(s) corresp. to  
 PT major histocompatibility complex antigens - esp. for delaying onset  
 PT of clinical symptoms of insulin dependent diabetes by modulating T  
 PT cell mediated attack on target cells  
 XX  
 PS Claim 7; Page 20; 24pp; English.  
 XX  
 CC AAW0512-W07518 represent T-cell modulating peptides that can be used in  
 CC the method of the invention. These sequences are based on a portion of  
 CC the generic peptide corresponding to residues 70-91 of the alpha-domain  
 CC of the major histocompatibility complex (MHC) class I antigen (see  
 CC AAW07510). The method is for affecting the course of an autoimmune  
 CC disease involving T-cell mediated destruction of tissue in mammals.  
 CC These peptides are used especially to treat insulin-dependent diabetes  
 CC mellitus, preferably being administered during the pre-clinical stage to  
 CC delay onset of the disease. Other diseases that can be treated are  
 CC multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus vulgaris,  
 CC Sjogren's disease, thyroid disease, Hashimoto's thyroiditis, myasthenia  
 CC gravis, etc. The peptides modulate T-cell mediated attack on autologous  
 CC target cells, and may also reduce inflammation, swelling, and release of  
 CC cytokines, perforins, granzymes etc. associated with T cell activation.  
 XX  
 SQ Sequence 10 AA;

Query Match 20.8%; Score 5; DB 17; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLL 23  
 |||||  
 Db 4 lrtll 8

RESULT 12  
 AAW33785  
 ID AAW33785 standard; peptide; 10 AA.  
 XX  
 AC AAW33785;  
 XX  
 DE 19-JUN-1998 (first entry)  
 XX  
 KM Peptide B2705.75-84 tested for immunomodulating activity.  
 XX  
 KM Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KM transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KM rejection.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9744351-A1.  
 XX

PD 27-NOV-1997.  
 XX  
 PF 22-MAY-1997; 97WO-US08689.  
 XX  
 PR 24-MAY-1996; 96US-0653294.  
 XX  
 PA (STRD) UNIV LELAND STANFORD JUNIOR.  
 XX  
 PI Buelow R, Clayberger C, Krensky AM;  
 XX  
 DR WPI: 1998-086530/08.  
 XX  
 PT New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 PT alpha-1 domain, used for preventing rejection of transplants or  
 PT treating autoimmune diseases  
 XX  
 PS Example 1; Page 19; 41pp; English.  
 XX  
 CC Peptides AAW33784-98 and AAW33778-9 were assayed for their  
 CC immunomodulating activity. A peptide-type compound or variant is claimed  
 CC which has immunomodulating activity, including the N-terminal acylated  
 CC and/or C-terminal amidated or esterified forms of up to 60 amino acids,  
 CC where the peptide-type compound comprises the formula: A-B, where A, B =  
 CC (R aa76-77/L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 =  
 CC D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or  
 CC small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino  
 CC acid. The sequence in the brackets may optionally be absent or truncated  
 CC at any peptide type bond within the brackets. The compounds comprise  
 CC amino acid sequences related to a Class I HLA-B alpha domain (positions  
 CC 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from  
 CC undesirably attacking cells in a host or in vitro. They can also be  
 CC used in combination with antigenic peptides or proteins of interest to  
 CC activate CTLs. They can also inhibit the proliferation of T cells in  
 CC response to anti-CD3. The peptide can be used for preventing rejection  
 CC of transplants or for treating autoimmune diseases, e.g. diabetes,  
 CC rheumatoid arthritis and lupus erythematosus. The products can also be  
 CC used for detection and diagnosis.  
 XX  
 SQ Sequence 10 AA;

Query Match 20.8%; Score 5; DB 19; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLL 23  
 |||||  
 Db 4 lrtll 8

RESULT 13  
 AAY72484  
 ID AAY72484 standard; peptide; 10 AA.  
 XX  
 AC AAY72484;  
 XX  
 DE 24-APR-2001 (first entry)  
 XX  
 KM Immunosuppressive peptide, DK.75-84, to prevent allograft rejection.  
 KM Immunosuppressive; allograft rejection; topological parameter;  
 KM physico-chemical parameter; in silico screening; pharmaceutical;  
 KM cosmetic; agrochemical; biomaterial; veterinary application.  
 XX  
 OS Unidentified.  
 OS  
 PN WO200079263-A2.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PE 15-MAY-2000; 2000WO-EP04338.  
 XX  
 PR 18-JUN-1999; 99EP-0401526.

XX (SYNT-) SYNT:EM SA.  
 PA Lahana R, Clair P, Yasri A;  
 XX WPI: 2001-091623/10.  
 DR  
 XX Identifying active candidate molecules on the basis of selected  
 PT physico-chemical parameters, for in silico screening of compounds  
 PT useful in pharmaceuticals, cosmetics, veterinary applications and  
 PT agrochemicals  
 PS Disclosure: Page 38; 55pp; English.  
 XX  
 CC The present invention relates to a method for identifying  
 CC physico-chemical and/or topological parameters associated with biological  
 CC activity. The method involves selecting the first subset from  
 CC predetermined set of physico-chemical parameters, determining their value  
 CC of function, and selecting the second subset from physico-chemical  
 CC parameters based on the values of function, such that each second subset  
 CC is more closely associated with the activity than the first subset. The  
 CC selected physico-chemical parameters are useful for developing criteria  
 CC for screening candidate molecules and are suitable for use in silico  
 CC screening of compounds. The compounds may be used in pharmaceuticals,  
 CC cosmetics, agrochemicals, biomaterials and in veterinary applications. It  
 CC is also useful as an antibiotic or antifungal agent. The present  
 CC sequence is an immunosuppressive peptide, DK-75-84, identified by in  
 CC silico screening. The immunosuppressive activity of the peptide that  
 CC prevents allograft rejection is tested in a heterotopic allograft model  
 CC of mouse.  
 CC  
 SQ Sequence 10 AA;

Query Match 20.8%; Score 5; DB 22; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLL 23  
 XX |||||  
 DB 4 lrtll 8

RESULT 14  
 AAY72488  
 ID AAY72488 standard; peptide; 10 AA.  
 XX  
 AC AAY72488;  
 XX  
 DT 24-APR-2001 (first entry)  
 XX  
 DE Immunosuppressive peptide, 2705.75-84, to prevent allograft rejection.  
 XX  
 KW Immunosuppressive; allograft rejection; topological parameter;  
 KW physico-chemical parameter; in silico screening; pharmaceutical;  
 KW cosmetic; agrochemical; biomaterial; veterinary application.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200079263-A2.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 15-MAY-2000; 2000MO-EP04338.  
 XX  
 PR 18-JUN-1999; 99EP-0401526.  
 XX  
 PA (SYNT-) SYNT:EM SA.  
 XX  
 PI Lahana R, Clair P, Yasri A;  
 XX  
 DR WPI: 2001-091623/10.  
 XX

PT Identifying active candidate molecules on the basis of selected  
 PT physico-chemical parameters, for in silico screening of compounds  
 PT useful in pharmaceuticals, cosmetics, veterinary applications and  
 PT agrochemicals  
 PS Disclosure: Page 38; 55pp; English.  
 XX  
 CC The present invention relates to a method for identifying  
 CC physico-chemical and/or topological parameters associated with biological  
 CC activity. The method involves selecting the first subset from  
 CC predetermined set of physico-chemical parameters, determining their value  
 CC of function, and selecting the second subset from physico-chemical  
 CC parameters based on the values of function, such that each second subset  
 CC is more closely associated with the activity than the first subset. The  
 CC selected physico-chemical parameters are useful for developing criteria  
 CC for screening candidate molecules and are suitable for use in silico  
 CC screening of compounds. The compounds may be used in pharmaceuticals,  
 CC cosmetics, agrochemicals, biomaterials and in veterinary applications. It  
 CC is also useful as an antibiotic or antifungal agent. The present  
 CC sequence is an immunosuppressive peptide, 2705.75-84, identified by in  
 CC silico screening. The immunosuppressive activity of the peptide that  
 CC prevents allograft rejection is tested in a heterotopic allograft model  
 CC of mouse.  
 CC  
 SQ Sequence 10 AA;

Query Match 20.8%; Score 5; DB 22; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLL 23  
 XX |||||  
 DB 4 lrtll 8

RESULT 15  
 AAW29421  
 ID AAW29421 standard; peptide; 13 AA.  
 XX  
 AC AAW29421;  
 XX  
 DT 25-FEB-1998 (first entry)  
 XX  
 DE Glucose transport enhancing peptide, generic formula.  
 XX  
 KW Glucose transport; major histocompatibility complex class I;  
 KW MHC; antigen; diabetes mellitus; diagnostic reagent; insulin receptor;  
 KW adrenergic receptor; IGF-I receptor; transferrin.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note="Ser is optionally modified by a peptide chain  
 FT of 1-12 natural aminoacid residues"  
 FT  
 FT Misc-difference 13  
 FT /note="This is a natural amino acid residue other than  
 FT Tyr, Phe, His or Trp; the D-form of a natural  
 FT amino acid residue other than Tyr, Phe, His or  
 FT Trp; -NR4-CHR3-COOH; or -NR4-CHR3-CH2OH; R3=H,  
 FT -R9-R10 or -NR6R7; R4=H or alkyl; or  
 FT R3+R4=alkylene; R6, R7=H, alkyl, cycloalkyl or  
 FT cycloalkylalkyl, or together form alkylene or  
 FT alkenylene; R9=single bond, alkyl, alkenyl,  
 FT alkoxy or lower aminoalkyl; R10=H; or cycloalkyl  
 FT or cycloalkenyl, both optionally substituted by  
 FT 1-5 substituents selected from 1-5 halo, 1-3 NO2,  
 FT 1-3 CN, 1-3 SH, 1-3 sulphonyl, 1-3 sulphonyl,  
 FT 1-3 sulphoxyl, 1-3 OH, 1-3 COOH, 1-3 haloalkyl,  
 FT 1-3 NR6R7, 1-3 alkylene-NR6R7 and 1 or 2 oxo;  
 FT R1=H; R2=OH"  
 XX

```

PM US516642-A.
XX 14-MAY-1996.
XX 16-NOV-1992; 92US-0976872.
XX 16-NOV-1992; 92US-0976872.
PR (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX PA
XX PI
XX Mapelll C, Meyers CA;
XX WPI: 1996-251003/25.
XX
XX New glucose transport enhancing poly:peptide compounds - derived from
PT major histocompatibility complex Class I antigen, used for treating
PT diabetes mellitus or as diagnostic reagent.
XX
XX Claim 1; Columns 11-12; 8pp; English.
XX
CC The present sequence represents the generic formula of new glucose
CC transport enhancing peptides, derived from the major histocompatibility
CC complex (MHC) Class I antigen. They can be used for treating diabetes
CC mellitus and in screening for compounds with a similar mode of
CC action. They enhance or reduce the physiological response of a
CC cell as a result of binding to certain cellular components, such as
CC receptors, transporters, receptor-bound ligands or other membrane-bound
CC components. Typically, the peptides (optionally covalently bound to
CC insulin or its functional segment) enhance glucose uptake in a
CC cell. They may also be used diagnostically as a ligand to determine the
CC presence of the cellular components described above. These peptides have
CC increased potency and stability in bioassay buffers compared with known
CC MHC Class I-derived peptides. They can be radiolabeled with
CC retention of activity, and are not subject to aggregation and gel
CC formation.
XX
SQ Sequence 13 AA;

```

```

Query Match          20.8%; Score 5; DB 17; Length 13;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 19 LRTLL 23
   |||||
Db 6 lrtll 10

```

Search completed: September 15, 2001, 12:54:50  
Job time: 109 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:53:21 ; Search time 12.52 seconds  
(without alignments)  
146.022 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_163  
Perfect score: 24  
Sequence: 1 HSKSEKLRERIKVCCEQLRTLLP 24

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4711

Minimum DB seq length: 0  
Maximum DB seq length: 24

Post-processing: Listing first 45 summaries

Database: PIR\_68:\*

1: p1r1:.\*  
2: p1r2:.\*  
3: p1r3:.\*  
4: p1r4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	16.7	6	2	B33932 Ig mu chain D regi
2	4	16.7	14	2	PC4382 dehydrin 4.5K poly
3	4	16.7	18	1	OHEC2 heat-stable entero
4	4	16.7	18	2	A60103 heat-stable entero
5	4	16.7	20	2	I38417 HLA-A11 - human (f
6	4	16.7	20	2	C49753 hypothetical prote
7	4	16.7	22	2	PM0431 leucyl aminopeptid
8	4	16.7	23	2	I48936 cryptid - mouse (
9	4	16.7	23	2	I49413 cryptid - western
10	4	16.7	23	2	S31210 collagen alpha 2(I
11	4	16.7	23	2	PH1641 Ig H chain V-D-J r
12	4	16.7	24	2	A84023 hypothetical prote
13	3	12.5	5	2	I40698 biotin B - Citroba
14	3	12.5	5	2	S69237 surface protein te
15	3	12.5	6	2	A11490 pyruvate kinase (E
16	3	12.5	7	2	PS0254 18K protein 5507 -
17	3	12.5	8	2	S16324 hypothetical prote
18	3	12.5	8	2	PM0323 Ig heavy chain CRD
19	3	12.5	8	2	PM0043 phosphatidylethano
20	3	12.5	9	2	S13889 phosphoenolpyruvat
21	3	12.5	9	2	A12872 transaldolase (EC
22	3	12.5	9	2	A11497 Ig heavy chain CDR
23	3	12.5	9	2	PS0231 24K protein 4302 -
24	3	12.5	10	2	PS0451 hemocyanin subunit
25	3	12.5	10	2	S61308 aryl hydrocarbon
26	3	12.5	10	2	S65715 protein OA30023 -
27	3	12.5	11	2	PT0081 protein-tyrosine k
28	3	12.5	11	2	C59151 wound-induced prot
29	3	12.5	11	2	S19775

30	3	12.5	11	2	S41747	chaperonin 10 homo
31	3	12.5	11	2	I54193	Rhesus blood group
32	3	12.5	11	2	PH0939	T-cell receptor be
33	3	12.5	11	2	A59146	conotoxin suba - c
34	3	12.5	11	2	B59146	conotoxin subb - c
35	3	12.5	12	1	UOGM2	wrotensin II - lon
36	3	12.5	12	2	S42765	wrotensin II - tel
37	3	12.5	12	2	JS0423	wrotensin II-A pep
38	3	12.5	12	2	S26558	T-cell receptor be
39	3	12.5	12	2	S01222	translation elonga
40	3	12.5	12	2	A58501	24K kidney and bla
41	3	12.5	12	2	T44420	hypothetical prote
42	3	12.5	12	2	JS0424	wrotensin II-B pep
43	3	12.5	12	4	S49073	fiame shifted cyta
44	3	12.5	13	1	NTKMAS	alpha-conotoxin SI
45	3	12.5	13	2	P00445	wrotensin II - lau

## ALIGNMENTS

RESULT 1  
B33932  
Ig mu chain D region (D23) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 16-Aug-1996  
C:Accession: B33932  
R:Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.  
Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989  
A:Title: Two murine natural polypeptide autoantibodies are encoded by nonmutated ger  
A:Reference number: A33932; MUID:89282823  
A:Accession: B33932  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-6 <BAC>  
A:Cross-references: GB:M27107  
C:Keywords: immunoglobulin

Query Match 16.7%; Score 4; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 EKLK 8  
DB 1 EKLK 4  
RESULT 2  
PC4382  
dehydrin 4.5K polypeptide - Soybean (fragment)  
N:Alternate names: acid soluble 26K protein  
C:Species: Glycine max cv. Nattoyoriyu (soybean)  
C:Date: 06-Nov-1997 #sequence\_revision 06-Nov-1997 #text\_change 07-May-1999  
C:Accession: PC4382  
R:Momma, M.; Harguichi, K.; Saito, M.; Chikuni, K.; Hatada, K.  
Biosect. Biotechnol. Biochem. 61, 1286-1291, 1997  
A:Title: Purification and characterization of the acid soluble 26-kDa polypeptide fro  
A:Reference number: PC4380; MUID:97446521  
A:Accession: PC4382  
A:Molecule type: protein  
A:Residues: 1-14 <OMM>  
A:Experimental source: seed  
C:Comment: This protein is rich in hydrophilic amino acid, and highly heat stable.  
Query Match 16.7%; Score 4; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 KEKL 7  
DB 5 KEKL 8

RESULT 3  
HEC2  
heat-stable enterotoxin ST-2 - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 06-Jul-1982 #sequence\_revision 06-Jul-1982 #text\_change 31-Dec-1996  
C:Accession: A01823  
R:Chan, S.K.; Giannella, R.A.  
J. Biol. Chem. 256, 7744-7746, 1981  
A:Title: Amino acid sequence of heat-stable enterotoxin produced by Escherichia coli pat  
A:Reference number: A01823; MUID:81264141  
A:Accession: A01823  
A:Molecule type: protein  
A:Residues: 1-18 <CHA>  
A:Experimental source: strain 18D, serotype 0.42:k86:H37  
C:Comment: This enterotoxin is one of several, of differing molecular sizes, produced by  
idues of the heat-stable enterotoxin ST-1.  
C:Superfamily: heat-stable enterotoxin ST  
C:Keywords: enterotoxin; heat-stable protein  
F:1-18/Product: heat-stable enterotoxin ST-2 #status experimental <MAT>  
F:5-10,6-14,9-17/Disulfide bonds: #status predicted

Query Match 16.7%; Score 4; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 YCCE 17  
||||  
DB 4 YCCE 7

RESULT 4  
A60103  
heat-stable enterotoxin ST-1a - Citrobacter freundii  
C:Species: Citrobacter freundii  
C:Date: 10-Nov-1992 #sequence\_revision 10-Nov-1992 #text\_change 18-Jun-1993  
C:Accession: A60103  
R:Guarino, A.; Giannella, R.; Thompson, M.R.  
Infect. Immun. 57, 649-652, 1989  
A:Title: Citrobacter freundii produces an 18-amino-acid heat-stable enterotoxin identical  
A:Reference number: A60103; MUID:89108617  
A:Accession: A60103  
A:Molecule type: protein  
A:Residues: 1-18 <GUA>  
C:Superfamily: heat-stable enterotoxin ST

Query Match 16.7%; Score 4; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 YCCE 17  
||||  
DB 4 YCCE 7

RESULT 5  
I38417  
HLA-A11 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 23-Jul-1999  
C:Accession: I38417  
R:Balas, A.; Garcia-Sanchez, F.; Gomez-Reino, F.; Vicario, J.L.  
Hum. Immunol. 41, 69-73, 1994  
A:Title: HLA class I allele (HLA-A2) expression defect associated with a mutation in its  
A:Reference number: I38417; MUID:95137784  
A:Accession: I38417  
A:Status: preliminary; translated from GR/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-20 <RES>  
A:Cross-references: EMBL:U02934; NID:9414543; PIDN:AAA76607.1; PID:9414544

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 16.7%; Score 4; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 RTLL 23  
||||  
DB 7 RTLL 10

RESULT 6  
C49753  
hypothetical protein (prob 5' region) - Serratia marcescens (fragment)  
C:Species: Serratia marcescens  
C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 08-Oct-1999  
C:Accession: C49753; S11643  
R:Omori, K.; Suzuki, S.I.; Imai, Y.; Komatsubara, S.  
J. Gen. Microbiol. 137, 509-517, 1991  
A:Title: Analysis of the Serratia marcescens proBA operon and feedback control of pro  
A:Reference number: A49753; MUID:91237315  
A:Accession: C49753  
A:Molecule type: DNA  
A:Residues: 1-20 <OMO>  
A:Cross-references: GB:X53086; NID:947251; PIDN:CAA37253.1; PID:947252

Query Match 16.7%; Score 4; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKL 7  
||||  
DB 12 KEKL 15

RESULT 7  
P70431  
leucyl aminopeptidase (EC 3.4.11.1) / prolyl aminopeptidase (EC 3.4.11.5) - human (fr  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 31-Mar-2000  
C:Accession: P70431; A42432  
R:Matsushima, M.; Takahashi, T.; Ichinose, M.; Miki, K.; Kurokawa, K.; Takahashi, K.  
Biochem. Biophys. Res. Commun. 178, 1459-1464, 1991  
A:Title: Structural and immunological evidence for the identity of prolyl aminopeptid  
A:Reference number: P70429; MUID:91337097  
A:Accession: P70431  
A:Molecule type: protein  
A:Residues: 1-22 <MAT>  
A:Experimental source: liver  
R:Haris, C.A.; Hunte, B.; Krauss, M.R.; Taylor, A.; Epstein, L.B.  
J. Biol. Chem. 267, 6865-6869, 1992  
A:Title: Induction of leucine aminopeptidase by interferon-gamma. Identification by p  
A:Reference number: A42432; MUID:92202241  
A:Accession: A42432  
A:Molecule type: protein  
A:Residues: 'xx', 3-10 <HAR>  
C:Comment: This enzyme catalyzes the removal of unsubstituted amino-terminal amino ac  
C:Superfamily: cytosol aminopeptidase  
C:Keywords: alpha-aminoacylpeptide hydrolase

Query Match 16.7%; Score 4; DB 2; Length 22;  
Best Local Similarity 100.0%; Pred. No. 7.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SKK 6  
||||  
DB 10 SKK 13

RESULT 8



I48936  
cryp1din - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 16-Jul-1999  
C:Accession: I48936  
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H.  
Mamm. Genome 5, 349-355, 1994  
A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.  
A:Reference number: I48934; MUID:94319082  
A:Accession: I48936  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-23 <RES>  
A:Cross-references: EMBL:U05705; NID:9497031; PIDN:AAB60468.1; PID:9497032  
C:Superfamily: mammalian defensin

Query Match  
Best Local Similarity 16.7%; Score 4; DB 2; Length 23;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 RRER 11  
|||  
Db 3 RRER 6

RESULT 9  
I49413  
cryp1din - western wild mouse (fragment)  
C:Species: Mus spretus (western wild mouse)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 16-Jul-1999  
C:Accession: I49413  
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H.  
Mamm. Genome 5, 349-355, 1994  
A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.  
A:Reference number: I48934; MUID:94319082  
A:Accession: I49413  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-23 <RES>  
A:Cross-references: EMBL:U05706; NID:9497033; PIDN:AAB60469.1; PID:9497034  
C:Superfamily: mammalian defensin

Query Match  
Best Local Similarity 16.7%; Score 4; DB 2; Length 23;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 RRER 11  
|||  
Db 3 RRER 6

RESULT 10  
S31210  
collagen alpha 2(I) chain precursor - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 13-Aug-1999  
C:Accession: S31210  
R:Guennet, D.R.; Ritzenthaler, J.D.; Foley, J.; Jackson, J.D.; Smith, B.D.  
Biochem. J. 283, 699-703, 1992  
A:Title: DNA methylation inhibits transcription of procollagen alpha-2(I) promoters.  
A:Reference number: S31210; MUID:92272666  
A:Accession: S31210  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-23 <GUE>  
A:Cross-references: EMBL:X66209; NID:955974; PIDN:CAA46960.1; PID:955975  
C:Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology  
C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix

Query Match 16.7%; Score 4; DB 2; Length 23;

Best Local Similarity 100.0%; Pred. No. 7.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 RTLL 23  
|||  
Db 8 RTLL 11

RESULT 11  
PH1641  
Ig H chain V-D-J region (clone B-lees 240) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 17-Mar-1999  
C:Accession: PH1641  
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.  
J. Exp. Med. 178, 317-329, 1993  
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-lees m  
A:Reference number: PH1580; MUID:93301609  
A:Accession: PH1641  
A:Molecule type: DNA  
A:Residues: 1-23 <LEV>  
A:Experimental source: bone marrow pre-B lymphocyte  
C:Keywords: immunoglobulin

Query Match  
Best Local Similarity 16.7%; Score 4; DB 2; Length 23;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LRRE 10  
|||  
Db 12 LRRE 15

RESULT 12  
A84023  
hypothetical protein BH2985 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: A84023  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: A83650; MUID:20263314  
A:Accession: A84023  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-24 <STO>  
A:Cross-references: GB:AP001517; GB:BA000004; NID:910175500; PIDN:BA06704.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH2985

Query Match  
Best Local Similarity 16.7%; Score 4; DB 2; Length 24;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LRRE 10  
|||  
Db 20 LRRE 23

RESULT 13  
I40698  
biotin B - Citrobacter freundii (fragment)  
C:Species: Citrobacter freundii  
C>Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 12-Aug-1996  
C:Accession: I40698  
R:Shiuan, D.; Campbell, A.  
Gene 67, 203-211, 1988  
A:Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citroba  
A:Reference number: I40697; MUID:89006280

Query Match 16.7%; Score 4; DB 2; Length 23;

A:Accession: I40698  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-5 <RES>  
A:Cross-references: GB:M21922; NID:g144434

Query Match 12.5%; Score 3; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 HSS 3  
|||  
DB 3 HSS 5

RESULT 14  
S69237  
surface protein tetrabrachion heavy chain - Staphylothermus marinus (fragment)  
C:Species: Staphylothermus marinus  
C:Date: 04-Dec-1997 #sequence\_revision 04-Dec-1997 #text\_change 17-Mar-1999  
C:Accession: S69237  
R:Peterson, J.; Mitsch, M.; Kuehlmoergen, B.; Golbik, R.; Lupas, A.; Kellermann, J.; Engel  
J. Mol. Biol. 245, 385-401, 1995  
A:title: Tetrabrachion: a filamentous archaebacterial surface protein assembly of unusua  
A:Reference number: S69237; M0ID:95139068  
A:Accession: S69237  
A:Molecule type: protein  
A:Residues: 1-5 <PEP>  
A:Experimental source: strain FL, DSM 3639  
C:Keywords: cell wall; glycoprotein; heat-stable protein

Query Match 12.5%; Score 3; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 TTL 23  
|||  
DB 2 TTL 4

RESULT 15  
A11490  
pyruvate kinase (EC 2.7.1.40) - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 03-Mar-1995  
C:Accession: A11490  
R:Hjeltnist, G.; Andersson, J.; Edlund, B.; Engstrom, L.  
Biochem. Biophys. Res. Commun. 61, 559-563, 1974  
A:title: Amino acid sequence of a (32-P)phosphopeptide from pig liver pyruvate kinase ph  
A:Reference number: A11490; M0ID:75127438  
A:Accession: A11490  
A:Molecule type: protein  
A:Residues: 1-6 <HLE>  
A:Experimental source: liver  
C:Keywords: glycolysis; phosphotransferase

Query Match 12.5%; Score 3; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LRR 9  
|||  
DB 1 LRR 3

Search completed: September 15, 2001, 12:55:13  
job time: 112 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:54:31 ; Search time 9.43 Seconds  
(without alignments)  
87.183 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_163  
Perfect score: 24  
Sequence: 1 HSKSEKLRERIRKCCQRLTLRP 24

Scoring table:  
OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1336

Minimum DB seq length: 0  
Maximum DB seq length: 24

Post-processing: Listing first 45 summaries

Database : SWISSPROT\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4	16.7	14	UC15_MAIZE	P80621 zea mays (m
2	4	16.7	15	UC27_MAIZE	P80633 zea mays (m
3	4	16.7	18	HSTR_ECOLI	P01560 escherichia
4	4	16.7	20	YPRB_SERMA	P22581 serratia ma
5	4	16.7	24	CT31_LITCI	P81851 litorea cit
6	4	16.7	24	FEDG_AMEYE	P80707 amycolatops
7	3	12.5	5	BIOB_CITFR	P12997 citropacter
8	3	12.5	9	TAL1_PICJA	P17441 pichia jadi
9	3	12.5	10	TAL3_PICJA	P17441 pichia jadi
10	3	12.5	10	TEMK_RANTE	P56923 rana tempor
11	3	12.5	12	KRI6_GINBI	P36207 ginkgo blio
12	3	12.5	12	TM2A_METMA	P80652 methanosarc
13	3	12.5	12	UR2A_CATCO	P04558 catostomus
14	3	12.5	12	UR2B_CATCO	P04559 catostomus
15	3	12.5	12	UR2B_CYPCA	P04561 cyprinus ca
16	3	12.5	12	UR2_GILMT	P01147 gillithys
17	3	12.5	12	UR2_POLSP	P81022 polyodon sp
18	3	12.5	13	UR2_SCYCA	P35490 scyllorhinu
19	3	12.5	13	CXAA_CONST	P28878 conus stria
20	3	12.5	13	CXET_CONTE	P81755 conus texti
21	3	12.5	13	IDHP_RAT	P56574 rattus norv
22	3	12.5	13	IRBP_MOUSE	P49194 mus musculu
23	3	12.5	13	TEMC_RANTE	P56918 rana tempor
24	3	12.5	13	TEMD_RANTE	P56919 rana tempor
25	3	12.5	14	CAT2_FASHE	P80342 fasciola he
26	3	12.5	15	EFIA_MICCR	P81266 micropilits
27	3	12.5	15	UC30_MAIZE	P80636 zea mays (m
28	3	12.5	16	IBP4_PIG	P24854 sus scrofa
29	3	12.5	17	EFG_THEAO	Q01697 thermus aqu
30	3	12.5	17	PC24_BRANU	P81097 brassica na
31	3	12.5	17	RANR_RANRU	P08952 rana rugosa
32	3	12.5	17	TPIS_PINPS	P81666 pinus pins
33	3	12.5	18	HEX_ADECU	P35985 canine aden

34	3	12.5	19	1	DHAB_COMTE	P80704 comamonas t
35	3	12.5	19	1	FIBB_VDUYU	P14482 vulpes vulp
36	3	12.5	19	1	LCRP_PETMA	Q10996 petromyzon
37	3	12.5	19	1	LPBG_ECOLI	P33236 escherichia
38	3	12.5	19	1	NUO6_SOLIU	P80729 solanum tub
39	3	12.5	19	1	UP28_UPEMT	P82040 uperoletia m
40	3	12.5	20	1	EPFU_MYCSY	P81407 mycoplasma
41	3	12.5	20	1	LTC_FELCA	P37155 felis silve
42	3	12.5	20	1	MLIT_BOVIN	P35451 bos taurus
43	3	12.5	20	1	YOHK_KLEAE	P56506 klebsiella
44	3	12.5	21	1	DIAG_AMEYE	P80414 amycolatops
45	3	12.5	21	1	FIBB_ANTFM	P14465 antilocapra

## ALIGNMENTS

RESULT 1						
ID	UC15_MAIZE	STANDARD:	PRT:	14 AA.		
AC	P80621;					
DT	01-OCT-1996 (Rel. 34, Created)					
DT	01-OCT-1996 (Rel. 34, Last sequence update)					
DT	15-JUL-1999 (Rel. 38, Last annotation update)					
DE	UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLOGICAL COLEOPTILE (SPOT 245)					
DE	(FRAGMENT)					
OS	zea mays (Maize).					
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;					
OC	Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;					
OC	Andropogoneae; Zea.					
OX	NCBI_TaxID=4577;					
RN	[1]					
RP	SEQUENCE.					
RP	TISSUE=Coleoptile;					
RA	Touzet P., Ricciardi F., Morin C., Damerval C., Huet J.-C.;					
RA	Pernollet J.-C., Zivy M., de Vienne D.;					
RT	"The maize two dimensional gel protein database: towards an integrated					
RT	genome analysis program."					
RL	Theor. Appl. Genet. 93:997-1005(1996).					
CC	-I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN					
CC	PROTEIN IS: 4.8, ITS MW IS: 35.7 KDA.					
DR	Maize-2DPAGE: P80621; COLEOPTILE.					
DR	MaizeDB: 123947; -.					
FT	NON-TER	1	14			
FT	NON-TER	1	14			
SQ	SEQUENCE	14 AA;	1396 MW;	CG8949275F404CD2 CRC64;		
Query Match						
		16.7%;	Score 4;	DB 1;	Length 14;	
		Best Local Similarity	100.0%;	Pred. No. 1.9e+02;		
		Matches 4;	Conservative	0;	Mismatches 0;	Indels 0;
QY	21 TLUP 24					
DB	3 TLUP 6					
RESULT 2						
ID	UC27_MAIZE	STANDARD:	PRT:	15 AA.		
AC	P80633;					
DT	01-OCT-1996 (Rel. 34, Created)					
DT	01-OCT-1996 (Rel. 34, Last sequence update)					
DT	15-JUL-1999 (Rel. 38, Last annotation update)					
DE	UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLOGICAL COLEOPTILE (SPOT 688)					
DE	(FRAGMENT)					
OS	zea mays (Maize).					
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;					
OC	Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;					
OC	Andropogoneae; Zea.					
OX	NCBI_TaxID=4577;					
RN	[1]					
RP	SEQUENCE.					

```

RC TISSUE-Coleoptile;
RA Toust P., Ricciardi F., Morin C., Damerval C., Huet J.-C.,
RA Perrotet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.4, ITS MW IS: 48.4 KDA.
CC -1- SIMILARITY: TO XENOPUS HISTONE-BINDING PROTEIN N1/N2 AND RABBIT
CC AND HUMAN NUCLEAR AUTOANTIGENIC SPERM PROTEIN.
DR Maize-2DPAGE: P80633; COLEOPTILE.
DR MaizeDB: 123958; -.
FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA: 1853 MW: CA0E12A5DAED8DC7 CRC64;

Query Match 16.7%; Score 4; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 EQLR 20
DB 3 EQLR 6

RESULT 3
HSTR_ECOLI STANDARD: PRT; 18 AA.
AC P01560:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HEAT-STABLE ENTEROTOXIN ST-2 (ST-B).
OS Escherichia coli.
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Escherichia.
OX NCBI_TaxID=562;
RX MEDLINE=87191003; PubMed=3552731;
RA Shimonishi Y., Hidaka Y., Koizumi M., Hane M., Almoto S., Takeda T.,
RA Miwatani T., Takeda Y.;
RT "Mode of disulfide bond formation of a heat-stable enterotoxin (STn)
RT produced by a human strain of enterotoxigenic Escherichia coli.";
RL FEBS Lett. 215:165-170(1987).
CC -1- FUNCTION: TOXIN WHICH ACTIVATES THE PARTICULATE FORM OF GUANYLATE
CC CYCLASE AND INCREASES CYCLIC GMP LEVELS WITHIN THE HOST
CC INTESTINAL EPITHELIAL CELLS.
CC -1- DISEASE: BOTH HEAT-STABLE AND HEAT-LABILE ENTEROTOXINS ARE
CC PRODUCED BY PATHOGENIC STRAINS OF E. COLI AND AFFECT THE DIGESTIVE
CC TRACT OF MAMMALS.
CC -1- SIMILARITY: BELONGS TO THE HEAT-STABLE ENTEROTOXIN FAMILY.
DR PIR: A01833; QHEC2.
DR HSSP: P01559; IETN.
DR InterPro: IPR001489; -.
DR PROSITE: PS00273; ENTEROTOXIN_H_STABLE; 1.
KW Toxin; Enterotoxin.
FT DISULFID 5 10
FT DISULFID 6 14
FT DISULFID 9 17
SQ SEQUENCE 18 AA: 1978 MW: D0C975F49D600650 CRC64;

Query Match 16.7%; Score 4; DB 1; Length 18;

```

```

Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 YCCE 17
DB 4 YCCE 7

RESULT 4
YPRB_SERMA STANDARD: PRT; 20 AA.
ID YPRB_SERMA
AC P22581:
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE HYPOTHETICAL PROTEIN IN PROB 5 REGION (FRAGMENT).
OS Serratia marcescens.
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Serratia.
OX NCBI_TaxID=615;
RX MEDLINE=91237315; PubMed=1851803;
RA Omori K., Suzuki S., Imai Y., Komatsubara S.;
RT "Analysis of the Serratia marcescens proBA operon and feedback
RT control of proline biosynthesis.";
RL J. Gen. Microbiol. 137:509-517(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL: D90351; BAA14363.1; -.
CC DR EMBL: X53086; CAA37253.1; -.
CC DR PIR: S11643; S11643.
CC DR PIR: C49753; C49753.
CC KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 20 AA: 2248 MW: 4DD7777735276674 CRC64;

Query Match 16.7%; Score 4; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KEKL 7
DB 12 KEKL 15

RESULT 5
CT31_LITCI STANDARD: PRT; 24 AA.
ID CT31_LITCI
AC P81851; P81852; P81853;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CITROPIN 3.1.2 (CONTAINS: CITROPIN 3.1.1; CITROPIN 3.1).
OS Litorea citropa (Australian blue mountains tree frog).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
CC Litorea.
OX NCBI_TaxID=94770;
RX MEDLINE=99435977; PubMed=10504394;
RA Wegener K.L., Wabnitz P.A., Carver J.A., Bowie J.H., Chia B.C.S.,

```

RA Wallace J.C., Tyler M.J.;  
 RT "Host defence peptides from the skin glands of the Australian blue  
 mountains tree-frog *Litoria citropa*. Solution structure of the  
 RT antibacterial peptide citropin 1.1.";  
 RL Eur. J. Biochem. 265:627-637(1999).  
 CC -1- TISSUE SPECIFICITY: DORSAL AND SUBVENTRAL SKIN GLANDS.  
 CC PEPTIDE 1 24 CITROPIN 3.1.2.  
 FT PEPTIDE 1 23 CITROPIN 3.1.1.  
 FT PEPTIDE 1 22 CITROPIN 3.1.  
 SQ SEQUENCE 24 AA; 2614 MW; C9001E295BDE15D CRC64;

Query Match  
 Best Local Similarity 16.7%; Score 4; DB 1; Length 24;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKL 7  
 DB 7 KEKL 10

RESULT 6  
 FEDG\_AMYME STANDARD; PRT; 24 AA.  
 ID FEDG\_AMYME  
 AC P80707;  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE FORMATE ESTER DEHYDROGENASE, GAMMA CHAIN (EC 1.2.99.-) (FEDH)  
 DE (FRAGMENT).  
 OS Amycolatopsis methanolica.  
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 CC Actinomycetales; Pseudonocardiales; Pseudonocardaceae;  
 CC Amycolatopsis.  
 CC NCBI\_TaxID=1814;  
 OX NCBI\_TaxID=1814;  
 RN [1]  
 RP STRAIN=NCIB 11946;  
 RC MEDLINE=96140591; PubMed=8554333;  
 RX Kim S.W., Luykx D.M.A.M., de Vries S., Duine J.A.;  
 RT "A second molybdoprotein aldehyde dehydrogenase from *Amycolatopsis*  
 RT methanolica NCIB 11946.";  
 RL Arch. Biochem. Biophys. 325:1-7(1996).  
 CC -1- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA AND A GAMMA  
 CC CHAIN.  
 CC Oxidoreductase.  
 KW NON\_TER 24  
 FT SEQUENCE 24 AA; 2746 MW; A93A8EA007D0FC6B CRC64;  
 SQ SEQUENCE 24 AA; 2746 MW; A93A8EA007D0FC6B CRC64;

Query Match  
 Best Local Similarity 16.7%; Score 4; DB 1; Length 24;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 RTLL 23  
 DB 19 RTLL 22

RESULT 7  
 BIOD\_CITPR STANDARD; PRT; 5 AA.  
 ID BIOD\_CITPR  
 AC P12997;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE BIOTIN SYNTHASE (EC 2.8.1.6) (BIOTIN SYNTHETASE) (FRAGMENT).  
 GN BIOD.  
 OS Citrobacter freundii.  
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Citrobacter.  
 OX NCBI\_TaxID=546;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=89006280; PubMed=2971595;  
 RA Shian D., Campbell A.;  
 RT "Transcriptional regulation and gene arrangement of *Escherichia coli*,  
 RT Citrobacter freundii and *Salmonella typhimurium* biotin operons.";  
 RL Gene 67:203-211(1988).  
 CC -1- CATALYTIC ACTIVITY: DEHYDROBIOTIN + (S) = BIOTIN.  
 CC -1- PATHWAY: LAST STEP IN BIOTIN BIOSYNTHESIS PATHWAY.  
 CC -1- SIMILARITY: BELONGS TO THE BIOTIN AND LIPIC ACID SYNTHETASES  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M21922; NOT ANNOTATED CDS.  
 KW Biotin biosynthesis; Iron-sulfur; Transferase.  
 FT NON\_TER 5  
 SQ SEQUENCE 5 AA; 532 MW; 75A5B1EDD6F00000 CRC64;

Query Match  
 Best Local Similarity 12.5%; Score 3; DB 1; Length 5;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HSS 3  
 DB 3 HSS 5

RESULT 8  
 TAIL\_PICJA STANDARD; PRT; 9 AA.  
 ID TAIL\_PICJA  
 AC P17440;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE TRANSALDOLASE I (EC 2.2.1.2) (FRAGMENT).  
 OS *Pichia jadinii* (Yeast) (Candida utilis).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; *Pichia*.  
 OX NCBI\_TaxID=4903;  
 RN [1]  
 RP MEDLINE=77110646; PubMed=556924;  
 RX Sun S.C., Joris L., Tsolas O.;  
 RT "Purification of crystallization of transaldolase isozyme I and  
 RT evidence for different genetic origin of isozymes I and III in  
 RT *Candida utilis*.";  
 RL Arch. Biochem. Biophys. 178:69-78(1977).  
 CC -1- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF  
 CC METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.  
 CC -1- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE  
 CC 3-PHOSPHATE = D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.  
 CC -1- PATHWAY: NONOXIDATIVE PART OF THE PENTOSE-PHOSPHATE PATHWAY.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY.  
 DR PIR: A12872; A12872.  
 DR InterPro: IPR001585; -  
 DR PROSITE: PS00958; TRANSALDOLASE\_2; PARTIAL.  
 DR PROSITE: PS01054; TRANSALDOLASE\_1; PARTIAL.  
 FT TRANSFERASE; Pentose shunt.  
 FT NON\_TER 1 9  
 FT NON\_TER 1 9  
 SQ SEQUENCE 9 AA; 1008 MW; 274F31AFOBEB1E058 CRC64;

Query Match  
 Best Local Similarity 12.5%; Score 3; DB 1; Length 9;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 21 TLL 23  
DB 7 TLL 9

## RESULT 9

TAL3\_PICJA STANDARD; PRT; 9 AA.  
ID TAL3\_PICJA  
AC P17441;  
DT 01-AUG-1990 (Rel. 15, last sequence update)  
DT 01-AUG-1990 (Rel. 15, last sequence update)  
DT 01-FEB-1994 (Rel. 28, last annotation update)  
DE TRANSALDOLASE III (EC 2.2.1.2) (FRAGMENT).  
OS Pichia jadinii (yeast) (Candida utilis).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Pichia.  
NCBI\_TaxID=4903;  
NM [1]  
RN  
RP  
RX MEDLINE=75145197; PubMed=1092268;  
RA Tsolas O.; Sun S.C.;  
RT "Isolation of a peptide containing a histidyl-cysteine sequence  
from the active center of transaldolase.";  
RL Arch. Biochem. Biophys. 167:525-533(1975).  
CC -1- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF  
METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.  
CC -1- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE  
3-PHOSPHATE = D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.  
CC -1- PATHWAY: NONOXIDATIVE PART OF THE PENTOSE-PHOSPHATE PATHWAY.  
CC -1- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY.  
DR PIR; A11497; A11497.  
DR InterPro; IPR001585; -.  
DR PROSITE; PS00958; TRANSALDOLASE\_2; PARTIAL.  
DR PROSITE; PS01054; TRANSALDOLASE\_1; PARTIAL.  
KW Transferrase; Pentose shunt.  
FT NON\_TER 1 1  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 1033 MW; 325A31A44EB1E058 CRC64;

Query Match 12.5%; Score 3; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+04; Mismatches 0; Indels 0; Gaps 0;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 21 TLL 23  
DB 7 TLL 9

RESULT 10  
TEMK\_RANTE STANDARD; PRT; 10 AA.  
ID TEMK\_RANTE  
AC P56923;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 01-OCT-2000 (Rel. 40, last sequence update)  
DT 01-OCT-2000 (Rel. 40, last annotation update)  
DE TEMPORIN K.  
OS Rana temporaria (European common frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.  
NCBI\_TaxID=8407;  
NM [1]  
RN  
RP  
RX MEDLINE=97175050; PubMed=9022710;  
RA Simmaco M.; Mignogna G.; Canofeni S.; Miele R.; Mangoni M.L.;  
RT "Temporins, antimicrobial peptides from the European red frog Rana  
temporaria.";  
RL Eur. J. Biochem. 242:788-792(1996).  
CC -1- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-POSITIVE  
BACTERIA.

CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -1- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN  
FAMILY.  
CC Amphibian skin; Antibiotic; Amidation; Multigene family.  
KW MOD\_RES 10 10  
FT SEQUENCE 10 AA; 1123 MW; 390549B337272457 CRC64;

Query Match 12.5%; Score 3; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03; Mismatches 0; Indels 0; Gaps 0;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 22 LRP 24  
DB 1 LRP 3

## RESULT 11

RL6\_GINBI STANDARD; PRT; 12 AA.  
ID RL6\_GINBI  
AC P36207;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, last sequence update)  
DT 01-FEB-1996 (Rel. 33, last annotation update)  
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S16 (FRAGMENT).  
GN RPS16.  
OS Ginkgo biloba (Ginkgo).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Ginkgoophyta; Ginkgoales; Ginkgoaceae; Ginkgo.  
NCBI\_TaxID=3311;  
NM [1]  
RN  
RP  
RX MEDLINE=95094313; PubMed=8001171;  
RA Richard M.; Tremblay C.; Bellemare G.;  
RT "Chloroplastic genomes of Ginkgo biloba and Chlamydomonas moewusii  
contain a chlB gene encoding one subunit of a light-independent  
protochlorophyllide reductase.";  
RL Curr. Genet. 26:159-165(1994).  
CC -1- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL; U01531; AAA66977.1; -.  
DR InterPro; IPR000307; -.  
DR PROSITE; PS00732; RIBOSOMAL\_S16; 1.  
KW Ribosomal protein; Chloroplast.  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1488 MW; 6700EDAF9D033734 CRC64;

Query Match 12.5%; Score 3; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03; Mismatches 0; Indels 0; Gaps 0;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 KLR 8  
DB 3 KLR 5

RESULT 12  
TM2A\_METMA STANDARD; PRT; 12 AA.  
ID TM2A\_METMA  
AC P80652;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, last sequence update)  
DT 01-OCT-1996 (Rel. 34, last annotation update)

DE ALTERNATIVE TETRAHYDROMETHANOPTERIN S-METHYLTRANSFERASE 28 KDA SUBUNIT  
 (EC 2.1.1.86) (N5-METHYLTETRAHYDROMETHANOPTERIN--COENZYME M  
 DE METHYLTRANSFERASE 28 KDA SUBUNIT) (FRAGMENT).  
 OS Methanosarcina mazei.  
 OC Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;  
 OC Methanosarcina.  
 OX NCBI\_TaxID=2209;  
 RN [1]

RP SEQUENCE.  
 RC STRAIN-DSM 3647 / GOEL;  
 RX MEDLINE=96370840; PubMed=8774736;  
 RA Lienard T., Becher B., Marschall M., Bowlen S., Gottschalk G.;  
 RT "Sodium ion translocation by N5-methyltetrahydromethanopterin:  
 coenzyme M methyltransferase from Methanosarcina mazei Go1  
 RT reconstituted in ether lipid liposomes.";  
 RL Eur. J. Biochem. 239:857-864(1996).  
 CC -1- FUNCTION: THIS ENZYME COMPLEX CATALYZES AN INTERMEDIATE STEP IN  
 METHANOGENESIS, THE FORMATION OF METHYL-COENZYME M AND  
 CC TETRAHYDROMETHANOPTERIN FROM COENZYME M AND N5-METHYL-  
 CC TETRAHYDROMETHANOPTERIN.  
 CC -1- CATALYTIC ACTIVITY: 5-METHYL-5,6,7,8-TETRAHYDROMETHANOPTERIN +  
 CC 2-(METHYLTETRAHYDROMETHANOPTERIN-5,6,7,8-TETRAHYDROMETHANOPTERIN +  
 CC -1- SUBUNIT: COMPOSED OF SIX DIFFERENT SUBUNITS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC Transferrase; Methyltransferase; Transmembrane; Methanogenesis.  
 KM NON\_TER  
 FT SEQUENCE 12 AA; 1321 MW; 6DE4A5766232D76B CRC64;  
 SQ

Query Match 12.5%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 EKL 7  
 |||  
 DB 2 EKL 4

RESULT 13  
 UR2A\_CATCO STANDARD; PRT; 12 AA.  
 AC P04558;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE UROTENSIN IIA (U-IIA).  
 OS Catostomus commersoni (White sucker).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Catostomidae; Catostomus.  
 OX NCBI\_TaxID=7971;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=84041959; PubMed=6138758;  
 RA McMaster D., Lederis K.;  
 RT "Isolation and amino acid sequence of two urotensin II peptides from  
 RT Catostomus commersoni urophyses.";  
 RL Peptides 4:367-373(1983).  
 CC -1- FUNCTION: UROTENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY  
 CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A  
 CC CORTICOTROPIN-RELEASING FACTOR.  
 CC -1- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.  
 DR PIR: J50423; J50423.  
 DR InterPro: IPR001483;  
 DR Pfam: PF02083; Urotensin\_II; 1.  
 DR PROSITE: PS00984; UROTENSIN\_II; 1.  
 KM Hormone.  
 FT DISULFID 6 11  
 SQ SEQUENCE 12 AA; 1336 MW; 969C76DBB879CEBA CRC64;

Query Match 12.5%; Score 3; DB 1; Length 12;

Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 KYC 15  
 |||  
 DB 9 KYC 11

RESULT 14  
 UR2B\_CATCO STANDARD; PRT; 12 AA.  
 AC P04559;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE UROTENSIN IIB (U-IIB).  
 OS Catostomus commersoni (White sucker).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Catostomidae; Catostomus.  
 OX NCBI\_TaxID=7971;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=84041959; PubMed=6138758;  
 RA McMaster D., Lederis K.;  
 RT "Isolation and amino acid sequence of two urotensin II peptides from  
 RT Catostomus commersoni urophyses.";  
 RL Peptides 4:367-373(1983).  
 CC -1- FUNCTION: UROTENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY  
 CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A  
 CC CORTICOTROPIN-RELEASING FACTOR.  
 CC -1- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.  
 DR PIR: J50424; J50424.  
 DR InterPro: IPR001483;  
 DR Pfam: PF02083; Urotensin\_II; 1.  
 DR PROSITE: PS00984; UROTENSIN\_II; 1.  
 KM Hormone.  
 FT DISULFID 6 11  
 SQ SEQUENCE 12 AA; 1437 MW; 73961BDBB879CEBB CRC64;

Query Match 12.5%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 KYC 15  
 |||  
 DB 9 KYC 11

RESULT 15  
 UR2B\_CYPCA STANDARD; PRT; 12 AA.  
 AC P04561;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE UROTENSIN II-BETA.  
 OS Cyprinus carpio (Common carp).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Cyprinidae; Cyprininae; Cyprinus.  
 OX NCBI\_TaxID=7962;  
 RN [1]  
 RP SEQUENCE.  
 RA Munekata E., Ohtaki T., Ichikawa T., McMaster D., Lederis K.;  
 RL (in) Rich D.H., Gross E. (eds.);  
 RL Proceedings of the 7th american peptide symposium, pp.69-72,  
 RL Pierce Chemical Co., Rockford IL. (1981).  
 CC -1- FUNCTION: UROTENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY  
 CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A  
 CC CORTICOTROPIN-RELEASING FACTOR.  
 CC -1- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.

DR InterPro: IPR001483; -  
DR Pfam: PF02083; Urotensin\_II; 1.  
DR PROSITE: PS00984; UROTENSIN\_II; 1.  
KW Hormone.  
FT DISULFID 6 11  
FT VARIANT 2 2 G -> S  
SQ SEQUENCE 12 AA; 1407 MW; 73960A9FB879CEBB CRC64;

Query Match 12.5%; Score 3; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 KYC 15  
111  
Db 9 KYC 11

Search completed: September 15, 2001, 12:56:17  
Job time: 106 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:54:11 : Search time 17.55 Seconds  
(without alignments)  
180.930 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_163  
Perfect score: 24  
Sequence: 1 HSSSEKLRRERIKYCEQLRTLLP 24

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6525

Minimum DB seq length: 0  
Maximum DB seq length: 24

Post-processing: Listing first 45 summaries

Database :

SPREMBL\_16:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	16.7	9	11 09QWG2	Q9qwg2 mus musculu
2	4	16.7	11	4 09UELO	Q9uel0 homo sapien
3	4	16.7	11	10 09S8X4	Q9s8x4 glycine max
4	4	16.7	12	14 08S631	Q8s631 avian retro
5	4	16.7	13	2 047693	Q47693 escherichia
6	4	16.7	14	11 0920G5	Q920g5 mus musculu
7	4	16.7	14	14 010229	Q10229 human immun
8	4	16.7	14	14 010230	Q10230 human immun
9	4	16.7	14	14 010235	Q10235 human immun
10	4	16.7	16	2 048417	Q48417 klebsiella
11	4	16.7	16	4 09UB15	Q9ub15 homo sapien
12	4	16.7	16	4 09UD21	Q9ud21 homo sapien
13	4	16.7	16	11 09JHM3	Q9jhm3 rattus norv
14	4	16.7	17	14 09OEX9	Q9oex9 human immun
15	4	16.7	17	14 09OEX8	Q9oex8 human immun
16	4	16.7	18	4 09UCF9	Q9ucf9 homo sapien
17	4	16.7	18	10 09S8P9	Q9s8p9 populus tre
18	4	16.7	18	13 09PRR7	Q9pr7 gallus gall
19	4	16.7	20	6 09TRI3	Q9tri3 bos taurus

20	4	16.7	20	7 019618	Q19618 homo sapien
21	4	16.7	23	4 016312	Q16312 homo sapien
22	4	16.7	23	13 P82397	P82397 litoria aur
23	4	16.7	23	13 P82398	P82398 litoria aur
24	4	16.7	23	13 P82400	P82400 litoria ran
25	4	16.7	24	2 005616	Q05616 staphylococ
26	4	16.7	24	2 09K8M1	Q9k8m1 bacillus ha
27	4	16.7	24	7 09PP00	Q9pp00 homo sapien
28	4	16.7	24	7 09TN29	Q9tn29 homo sapien
29	4	16.7	24	7 09MT61	Q9mt61 homo sapien
30	4	16.7	24	7 09G1Z8	Q9g1z8 homo sapien
31	4	16.7	24	11 09QV68	Q9qv68 rattus sp.
32	4	12.5	7	14 066205	Q66205 porcine tra
33	3	12.5	8	5 09M6M5	Q9m6m5 toxoplasma
34	3	12.5	8	6 018854	Q18854 canis fam11
35	3	12.5	9	2 099193	Q99193 pseudomonas
36	3	12.5	9	2 047063	Q47063 escherichia
37	3	12.5	9	4 095574	Q95574 homo sapien
38	3	12.5	9	4 09UE09	Q9ue09 homo sapien
39	3	12.5	9	4 09UE26	Q9ue26 homo sapien
40	3	12.5	9	5 09TWD6	Q9twd6 leptinotars
41	3	12.5	9	6 09GK05	Q9gk05 capra hircu
42	3	12.5	9	6 09GK02	Q9gk02 capra hircu
43	3	12.5	10	2 051812	Q51812 escherichia
44	3	12.5	10	8 09MJ05	Q9mj05 podospora c
45	3	12.5	10	11 008622	Q08622 rattus norv

#### ALIGNMENTS

RESULT 1  
Q9QWG2 ID Q9QWG2 PRELIMINARY: PRT: 9 AA.  
AC Q9QWG2:  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE PROTEIN G BETA-2 SUBUNIT (FRAGMENT).  
GN GNB2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98030528; PubMed=9365246;  
RA Chretien S.; Duprez V.; Maouche L.; Gisselbrecht S.; Mayeux P.,  
RA Lacombe C.;  
RT "Abnormal erythropoietin (Epo) gene expression in the murine  
RT erythroleukemia IW32 cells is issued from a rearrangement between the  
RT G-protein b2 gene subunit and the Epo genes.";  
RL Oncogene 15:1995-1999(1997).  
DR EMBL; Y11970; CAA72706.1; -;  
FT NON-TER  
SQ SEQUENCE 9 AA; 1133 MW; 86FD0736DB172B05 CRC64:

Query Match 16.7%; Score 4; DB 11; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 EQLR 20  
DB 5 EQLR 8

RESULT 2  
ID Q9UELO PRELIMINARY: PRT: 11 AA.  
AC Q9UELO:  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE FAS ANTIGEN (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BLOOD;  
 RX MEDLINE=95355401; PubMed=7543095;  
 RA Wada N., Matsumura M., Ohba Y., Kobayashi N., Takizawa T.,  
 RA Nakaniishi Y.,  
 RT "Transcription stimulation of the Fas-encoding gene by nuclear factor  
 RT for interleukin-6 expression upon influenza virus infection.";  
 RL J. Biol. Chem. 270:18007-18012(1995).  
 DR EMBL; D31968; BAA20850.1; -;  
 FT NON-TER  
 SQ SEQUENCE 11 AA; 1256 MW; 0802145E2731B370 CRC64;

Query Match 16.7%; Score 4; DB 4; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 TLLP 24  
 ||||  
 Db 6 TLLP 9

RESULT 3  
 OQS8X4 PRELIMINARY; PRT; 11 AA.  
 AC OQS8X4;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE VEGETATIVE STORAGE PROTEIN 94 PEPTIDE 3, VS94-LIPOXYGENASE.  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 OC Fabales; Fabaceae; Papilionoideae; Glycine.  
 OX NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92361246; PubMed=1822994;  
 RA Tranbarger T.J., Franceschi V.R., Hildebrand D.F., Grimes H.D.;  
 RT "The soybean 94-kilodalton vegetative storage protein is a  
 RT lipoxigenase that is localized in paraveinal mesophyll cell  
 RT vacuoles.";  
 RL Plant Cell 3:973-987(1991).  
 SO SEQUENCE 11 AA; 1366 MW; 9B337C3C0DD9CB1A CRC64;

Query Match 16.7%; Score 4; DB 10; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 RERI 12  
 ||||  
 Db 8 RERI 11

RESULT 4  
 O85631 PRELIMINARY; PRT; 12 AA.  
 AC O85631;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE PROVIRAL DNA, MYC TO 3' LTR (FRAGMENT).  
 OS Avian retrovirus MH2.  
 OC Viruses; Retroid viruses; Retroviridae; Avian type C retroviruses.  
 OX NCBI\_TaxID=11870;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8503920; PubMed=6092695;  
 RA Suttrave P., Jensen H.W., Bister K., Rapp U.R.;  
 RT "3'-Terminal region of avian carcinoma virus MH2 shares sequence  
 RT elements with avian sarcoma viruses Y73 and SR-A.";  
 RL J. Virol. 52:703-705(1984).  
 DR EMBL; K03100; AAA42388.1; -;  
 FT NON-TER  
 SQ SEQUENCE 12 AA; 1466 MW; 72E4B884F30736DB CRC64;

Query Match 16.7%; Score 4; DB 14; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 EQLR 20  
 ||||  
 Db 5 EQLR 8

RESULT 5  
 O47693 PRELIMINARY; PRT; 13 AA.  
 AC O47693;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 GN TURB.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82080657; PubMed=7312036;  
 RA Hudson L., Rossi J., Landy A.;  
 RT "Dual function transcripts specifying tRNA and mRNA.";  
 RL Nature 294:422-427(1981).  
 DR EMBL; X04181; CAA27777.1; -;  
 KW Elongation factor; Protein biosynthesis.  
 FT NON-TER  
 SQ SEQUENCE 13 AA; 1617 MW; C433BE82A18D0B19 CRC64;

Query Match 16.7%; Score 4; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SKEK 6  
 ||||  
 Db 2 SKEK 5

RESULT 6  
 O920G5 PRELIMINARY; PRT; 14 AA.  
 AC O920G5;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE SHORTER ISOFORM OF INTERLEUKIN 15 (FRAGMENT).  
 GN INTERLEUKIN 15.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALBC;  
 RA Nishimura H.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

[2]  
RN SEQUENCE FROM N.A.  
RC STRAIN-BALBC;  
RX MEDLINE-98211658; PubMed-9551932;  
RA Nishimura H., Washizu J., Nakamura N., Enomoto A., Yoshikai Y.;  
RT "Translational efficiency is up-regulated by alternative exon in  
RL murine IL-15 mRNA."  
RL J. Immunol. 160:936-942(1998).  
RN  
RN [3]  
RN SEQUENCE FROM N.A.  
RA Prinz M., Hantsch U.K., Kettenmann H., Kirchhoff F.;  
RT "Alternative splicing of mouse IL-15 is due to the use of an internal  
RT splice site in exon 5."  
RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AB023307; BAA37122.1; -;  
DR EMBL: AC012587; CAI10069.1; -;  
FT NON\_TER 14 14  
SQ SEQUENCE 14 AA; 1503 MW; 62A43C388924E2B CRC64;

Query Match 16.7%; Score 4; DB 11; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSKE 5  
DB 4 SSKE 7

RESULT 7  
ID 010229 PRELIMINARY; PRT; 14 AA.  
AC 010229;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-98216723; PubMed-9557645;  
RA Salvi R., Garbuglia A.R., Di Caro A., Pulciani S., Montella F.,  
RA Benedetto A.;  
RT "Grossly defective nef gene sequences in a human immunodeficiency  
RT virus type 1-seropositive long-term nonprogressor.";  
RL J. Virol. 72:3646-3657(1998).  
DR EMBL: U89849; AAC26088.1; -;  
DR InterPro: IPR000328; -;  
DR Pfam: PF00517; GP41; 1.  
KW Envelope protein.  
FT NON\_TER 1 1  
SQ SEQUENCE 14 AA; 1721 MW; D5E75B69D45D0566 CRC64;

Query Match 16.7%; Score 4; DB 14; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 RTLL 23  
DB 11 RTLL 14

RESULT 8  
ID 010230 PRELIMINARY; PRT; 14 AA.  
AC 010230;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE ENVELOPE GLYCOPRO  
GN ENV.  
OS Human immunodeficiency  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-98216723; PubMed-9557645;  
RA Salvi R., Garbuglia A.R., Di Caro A., Pulciani S., Montella F.,  
RA Benedetto A.;  
RT "Grossly defective nef gene sequences in a human immunodeficiency  
RT virus type 1-seropositive long-term nonprogressor.";  
RL J. Virol. 72:3646-3657(1998).  
DR EMBL: U89850; AAC26094.1; -;  
DR InterPro: IPR000328; -;  
DR Pfam: PF00517; GP41; 1.  
KW Envelope protein.  
FT NON\_TER 1 1  
SQ SEQUENCE 14 AA; 1721 MW; D5E75B69D45D0566 CRC64;

Query Match 16.7%; Score 4; DB 11; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 RTLL 23  
DB 11 RTLL 14

RESULT 9  
ID 010235 PRELIMINARY; PRT; 14 AA.  
AC 010235;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-98216723; PubMed-9557645;  
RA Salvi R., Garbuglia A.R., Di Caro A., Pulciani S., Montella F.,  
RA Benedetto A.;  
RT "Grossly defective nef gene sequences in a human immunodeficiency  
RT virus type 1-seropositive long-term nonprogressor.";  
RL J. Virol. 72:3646-3657(1998).  
DR EMBL: U89855; AAC26094.1; -;  
DR InterPro: IPR000328; -;  
DR Pfam: PF00517; GP41; 1.  
KW Envelope protein.  
FT NON\_TER 1 1  
SQ SEQUENCE 14 AA; 1721 MW; D5E75B69D45D0566 CRC64;

Query Match 16.7%; Score 4; DB 14; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 RTLL 23  
DB 11 RTLL 14

RESULT 10  
ID 048417 PRELIMINARY; PRT; 16 AA.  
AC 048417;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-NOV-1998 (TREMBLrel. 08, last annotation update)  
 DE REPB.  
 GN REPB.  
 OS Klebsiella pneumoniae.  
 OC Plasmid pMT407.1.  
 DE Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Klebsiella.  
 OX NCBI\_TaxId=573;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95370152; PubMed=7543895;  
 RA Athanasopoulos V., Praszkie J., Pittard A.J.;  
 RT "The replication of an Incl/M plasmid is subject to antisense  
 control."  
 RT J. Bacteriol. 177:4730-4741(1995).  
 DR EMBL: U27345; AAA87027.1; -  
 KW Plasmid.  
 SO SEQUENCE 16 AA; 1927 MW; 14D95D0594A80E02 CRC64;

Query Match 16.7%; Score 4; DB 2; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 RTLL 23  
 DB 2 RTLL 5

RESULT 11  
 O9UB15 PRELIMINARY; PRT; 16 AA.  
 AC 09UB15;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)  
 DE CYSTATIN S.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92138674; PubMed=1778989;  
 RA Isemura S., Saitoh E., Sanada K., Minakata K.;  
 RT "Identification of full-sized forms of salivary (S-type) cystatins  
 (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of  
 cystatin S) in human whole saliva and determination of phosphorylation  
 sites of cystatin S."  
 RT J. Biochem. 110:648-654(1991).  
 RN [2]  
 RP SEQUENCE.  
 RX MEDLINE=92074898; PubMed=1741693;  
 RA Johnson M., Richardson C.F., Bergey E.J., Levine M.J.,  
 RA Nancollas G.H.;  
 RT "The effects of human salivary cystatins and statherin on  
 RT hydroxyapatite crystallization."  
 RT Arch. Oral Biol. 36:631-636(1991).  
 SO SEQUENCE 16 AA; 1765 MW; A6EDF31C3186F9C9 CRC64;

Query Match 16.7%; Score 4; DB 4; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSKS 5  
 DB 2 SSKS 5

RESULT 12  
 O9UD21 PRELIMINARY; PRT; 16 AA  
 ID O9UD21

AC O9UD21;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)  
 DE CYCLIN E-L (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95257942; PubMed=7739542;  
 RA Ohtsubo M., Theodoras A.M., Schumacher J., Roberts J.M., Pagano M.;  
 RT "Human cyclin E, a nuclear protein essential for the G1-to-S phase  
 transition."  
 RT Mol. Cell. Biol. 15:2612-2624(1995).  
 SO SEQUENCE 16 AA; 2089 MW; 777EFC69C445E29C CRC64;

Query Match 16.7%; Score 4; DB 4; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 RPR 11  
 DB 5 RPR 8

RESULT 13  
 O9JHM3 PRELIMINARY; PRT; 16 AA.  
 AC 09JHM3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)  
 DE MHC CLASS I ANTIGEN (FRAGMENT).  
 GN RT1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus  
 OX NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LEW;  
 RA Lambracht-Washington D., Fischer Lindahl K., Monigelt K.;  
 RT "What can class I promoter sequences tell us about evolution,  
 RT rat MHC?"  
 RT Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF210330; AAF74411.1; -  
 DR InterPro: IPR001039; -  
 DR Prodom: PD000050; -; 1.  
 FT NON\_TER 16  
 SO SEQUENCE 16 AA; 1635 MW; 647FF5119E02977D CRC64;

Query Match 16.7%; Score 4; DB 1;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatch 0;

OY 20 RTLL 23  
 DB 4 RTLL 7

RESULT 14  
 O9QEX9 PRELIMINARY; PRT; 16 AA.  
 AC 09QEX9;  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 14, last annotation update)  
 DE NEF PROTEIN (FRAGMENT)  
 GN NEF

OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lin H.J., Siwak E.B., Hollinger F.B.;  
 RT "Mutation rate of human immunodeficiency virus type 1 genomic RNA  
 RT deduced from long term culture of its biological clones.";  
 RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF178661; AAF04367.1; -;  
 DR InterPro; IPR000081; -;  
 DR ProDom; PD001306; -; 1.  
 FT NON\_TER 1  
 FT NON\_TER 17  
 SQ SEQUENCE 17 AA; 1960 MW; 9315C3C6F3515653 CRC64;

Query Match 16.7%; Score 4; DB 14; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 RERI 12  
 ||||  
 Db 14 RERI 17

RESULT 15  
 O9QEX8 PRELIMINARY; PRT; 17 AA.  
 AC O9QEX8;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE NEF PROTEIN (FRAGMENT).  
 GN NEF.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lin H.J., Siwak E.B., Hollinger F.B.;  
 RT "Mutation rate of human immunodeficiency virus type 1 genomic RNA  
 RT deduced from long term culture of its biological clones.";  
 RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF178662; AAF04368.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 17  
 SQ SEQUENCE 17 AA; 2032 MW; 919FC3C6F3515653 CRC64;

Query Match 16.7%; Score 4; DB 14; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 RERI 12  
 ||||  
 Db 14 RERI 17

Search completed: September 15, 2001, 12:55:59  
 Job time: 108 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:47:10 ; Search time 65.58 Seconds  
(without alignments)  
22.186 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_163

Perfect score: 24

Sequence: 1 HSKSEKLRERIKYCCQLRLTLP 24

Scoring table:

OLIGO  
Gapop 60.0 , Capext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : A.Geneseq\_0601.\*

```
1: /cgnl_9/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /cgnl_9/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /cgnl_9/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /cgnl_9/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /cgnl_9/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /cgnl_9/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /cgnl_9/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /cgnl_9/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /cgnl_9/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /cgnl_9/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /cgnl_9/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /cgnl_9/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /cgnl_9/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /cgnl_9/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /cgnl_9/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /cgnl_9/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /cgnl_9/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /cgnl_9/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /cgnl_9/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /cgnl_9/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /cgnl_9/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /cgnl_9/gcgdata/geneseq/geneseq/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	100.0	405	21	AAV79269 Human testis-speci
2	14	58.3	15	21	AAV79270 PHELIIX peptide use
3	8	33.3	1220	17	AAW03664 I2C-1 protein conf
4	8	33.3	1240	17	AAW03665 I2C-2 protein conf
5	8	33.3	1266	18	AAW25157 Tomato immunity 2
6	8	33.3	1266	21	AAW07754 Amino acid sequenc
7	7	29.2	135	21	AAW43791 Arabidopsis thalia
8	7	29.2	250	21	AAW43790 Arabidopsis thalia
9	7	29.2	291	21	AAW15758 Arabidopsis thalia
10	7	29.2	341	21	AAW15757 Arabidopsis thalia
11	7	29.2	443	21	AAW15756 Arabidopsis thalia

12	6	25.0	150	17	AAW02605 Human leukotriene
13	6	25.0	150	17	AAW02605 Human leukotriene C
14	6	25.0	152	21	AAW54945 Arabidopsis thalia
15	6	25.0	220	21	AAW32041 Arabidopsis thalia
16	6	25.0	236	21	AAW32040 Arabidopsis thalia
17	6	25.0	247	21	AAW30797 Arabidopsis thalia
18	6	25.0	248	21	AAW30796 Arabidopsis thalia
19	6	25.0	250	21	AAW42198 Human ORX ORF1962
20	6	25.0	252	21	AAW39232 Arabidopsis thalia
21	6	25.0	253	21	AAW39231 Arabidopsis thalia
22	6	25.0	259	21	AAW05356 Arabidopsis thalia
23	6	25.0	259	21	AAW45159 Arabidopsis thalia
24	6	25.0	260	21	AAW05355 Arabidopsis thalia
25	6	25.0	260	21	AAW45158 Arabidopsis thalia
26	6	25.0	268	20	AAW92839 E. coli dsbg prote
27	6	25.0	271	21	AAW20985 Arabidopsis thalia
28	6	25.0	284	21	AAW22161 Arabidopsis thalia
29	6	25.0	284	21	AAW32039 Arabidopsis thalia
30	6	25.0	291	21	AAW18003 Arabidopsis thalia
31	6	25.0	291	21	AAW49791 Arabidopsis thalia
32	6	25.0	293	19	AAW69419 Protein encoded by
33	6	25.0	297	21	AAW30795 Arabidopsis thalia
34	6	25.0	298	21	AAW20984 Arabidopsis thalia
35	6	25.0	302	21	AAW39230 Arabidopsis thalia
36	6	25.0	305	21	AAW56616 Human prostate can
37	6	25.0	310	21	AAW05354 Arabidopsis thalia
38	6	25.0	310	21	AAW45157 Arabidopsis thalia
39	6	25.0	311	21	AAW49790 Arabidopsis thalia
40	6	25.0	316	21	AAW16294 Eucalyptus grandis
41	6	25.0	316	21	AAW16321 Pinus radiata ane
42	6	25.0	316	22	AAW65730 Annexin-like prote
43	6	25.0	320	21	AAW18002 Arabidopsis thalia
44	6	25.0	329	21	AAW22160 Arabidopsis thalia
45	6	25.0	333	21	AAW22159 Arabidopsis thalia

#### ALIGNMENTS

RESULT 1	
AAV79269	
ID	AAV79269 standard; Protein; 405 AA.
XX	
AC	AAV79269;
XX	
DT	03-JUL-2000 (first entry)
XX	
DE	Human testis-specific transcription factor PHELIIX.
XX	
KW	PHELIIX; human; testis-specific; transcription factor;
KW	prostate cancer; bladder cancer; ovary cancer; testicular cancer;
KW	therapy; diagnosis; vaccine.
XX	
OS	Homo sapiens.
XX	
FH	
FT	Key
FT	Peptide
FT	Location/Qualifiers
FT	/note="nuclear localization signal"
FT	134..150
FT	Peptide
FT	/note="nuclear localization signal"
FT	163..169
FT	Domain
FT	140..189
FT	/note="basic Helix-Loop-Helix domain"
XX	
PN	WO200012709-A2.
XX	
PD	09-MAR-2000.
XX	
PF	31-AUG-1999; 99WO-US20137.
XX	
PR	31-AUG-1998; 98US-0098610.
PR	31-OCT-1998; 98US-0106524.
XX	
PA	(UROC-) UROGENESIS INC.

PA (AFAR/) AFAR D E.  
 PA (HUBE/) HUBERT R S.  
 PA (RAIT/) RAITANO A B.  
 PI Afar DE, Hubert RS, Raitano AB;  
 XX  
 DR MPI: 2000-237872/20.  
 DR N-PSDB; AA294275.  
 XX  
 PT Testis specific Helix loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -  
 XX  
 PS Claim 1; Fig 2A-D; 62pp; English.  
 XX  
 CC This sequence is that of human PHELIIX, a novel basic Helix loop  
 CC Helix protein thought to act as a transcription factor. PHELIIX  
 CC normally exhibits a testis-specific expression pattern but is  
 CC up-regulated in prostate and other types of cancer. The invention  
 CC provides diagnostic and therapeutic methods useful in the  
 CC management of various cancers which express PHELIIX, including  
 CC prostate cancer, bladder cancer, ovarian cancer and testicular  
 CC cancer, including therapies aimed at inhibition the transcription,  
 CC translation, processing or function of PHELIIX. The expression  
 CC pattern of PHELIIX suggests that is an ideal target for a cancer  
 CC vaccine approach to prostate cancer. PHELIIX protein can also be  
 CC used to screen for agonists and antagonists of therapeutic value  
 CC and to raise antibodies.  
 CC  
 SQ Sequence 405 AA;

Query Match 100.0%; Score 24; DB 21; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-17;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HSKSEKLRREIRIKYCCQELRTILP 24  
 DB 140 hsksekIrreIrIkYccqELrTIlp 163

RESULT 2  
 AAY79270  
 ID AAY79270 standard; Peptide; 15 AA.  
 XX  
 AC AAY79270;  
 XX  
 DT 03-JUL-2000 (first entry)  
 XX  
 DE PHELIIX peptide used to raise antibody.  
 XX  
 KW PHELIIX: human; testis-specific; transcription factor;  
 KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
 KW therapy; diagnosis; vaccine; antibody.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO200012709-A2.  
 PN  
 PD 09-MAR-2000.  
 PD  
 XX 31-AUG-1999; 99WO-US20137.  
 XX  
 PF 31-AUG-1998; 98US-0098610.  
 PR 31-OCT-1998; 98US-0106524.  
 PR  
 XX (UROC-) UROGENESIS INC.  
 PA (AFAR/) AFAR D E.  
 PA (HUBE/) HUBERT R S.  
 PA (RAIT/) RAITANO A B.  
 XX  
 PI Afar DE, Hubert RS, Raitano AB;  
 XX

DR MPI: 2000-237872/20.  
 XX  
 PT Testis specific Helix loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -  
 XX  
 PS Example 5; Page 36; 62pp; English.  
 XX  
 CC The present sequence is that of a peptide derived from human  
 CC PHELIIX (see AAY79269), a novel transcription factor that is  
 CC normally expressed only in testis tissue, but which is up-regulated  
 CC in prostate and some other cancers. The peptide was conjugated to  
 CC keyhole limpet haemocyanin and used to raise polyclonal antiserum  
 CC in rabbit. The antiserum demonstrated specificity for PHELIIX and  
 CC may therefore be useful for assessing the expression of PHELIIX in  
 CC patient samples.  
 CC  
 SQ Sequence 15 AA;

Query Match 58.3%; Score 14; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 7e-08;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HSKSEKLRREIRIKY 14  
 DB 1 hsksekIrreIrIkY 14

RESULT 3  
 AAW03664  
 ID AAW03664 standard; Protein; 1220 AA.  
 XX  
 AC AAW03664;  
 XX  
 DT 22-FEB-1997 (first entry)  
 XX  
 DE I2C-1 protein conferring Fusarium wilt disease-resistance.  
 XX  
 KW Tomato; I2C-1; Fusarium wilt; disease-resistance; Fusarium oxysporum;  
 KW P-loop; leucine zipper; leucine-rich repeat; transgenic plant;  
 KW restriction fragment length polymorphism; screening;  
 KW crop improvement; Solanaceae.  
 XX  
 OS Lycopersicon esculentum.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Region 66..71 "Repeat sequence"  
 FT Region 103..108 "Repeat sequence"  
 FT Region 200..207 "Repeat sequence"  
 FT Domain 256..267 "P-loop"  
 FT Region 256..267 "Repeat sequence"  
 FT Region 269..280 "Repeat sequence"  
 FT Region 286..295 "Repeat sequence"  
 FT Domain 317..324 "Conserved motif of unknown function"  
 FT Domain 380..385 "Conserved motif of unknown function"  
 FT Domain 400..401 "Conserved motif of unknown function"  
 FT Misc-difference 419..430 "Conserved motif of unknown function"  
 FT Domain 498..506 "Conserved motif of unknown function"  
 FT Domain 559..623 "Conserved motif of unknown function"  
 FT Domain 624..645 "Leucine-rich repeat region"  
 FT Domain "note="Putative leucine zipper domain"



FT Domain 646..1220  
 FT /note= "Leucine-rich repeat region"  
 XX  
 PM W09632007-A1.  
 PD 17-OCT-1996.  
 XX  
 PF 15-APR-1996; 96WO-US05272.  
 XX  
 PR 13-APR-1995; 95IL-0113373.  
 PA (YEDA ) YEDA RES & DEV CO LTD.  
 PA (YISS ) YISSUM RES & DEV CO.  
 XX  
 PI Eshed Y, Fluhr R, Ori N, Paran I, Zamir D;  
 DR WPI: 1996-476739/47.  
 DR N-PSDB; AAT42134.  
 XX  
 PT New DNA conferring resistance to Fusarium sp. - used for producing  
 PT disease-resistant tomato plants and for identifying new resistance  
 PT genes and diseases  
 PS  
 PS Claim 1; Fig 4A-C; 50pp; English.  
 XX  
 CC This I2C-1 protein is encoded by a sequence from the I2C multigene  
 CC family from the I2 Fusarium wilt resistance locus of tomato, and  
 CC confers resistance to Fusarium oxysporum f.sp. Lycopersici race-2.  
 CC This sequence and I2C-2 (AAW03665) are encoded by genes from a locus  
 CC completely linked to I2, and show structural similarity with other  
 CC resistance proteins. The protein has a conserved N-terminal  
 CC nucleotide-binding domain (the P-loop) and 5 other conserved  
 CC domains of unknown function. At least half the C-terminus is  
 CC composed of leucine-rich repeats, which may be responsible for  
 CC specificity of interaction, either with a pathogen protein  
 CC component, or with downstream factors involved with signal  
 CC transduction. There does not appear to be a transmembrane domain,  
 CC indicating an intracellular location. A putative leucine zipper  
 CC domain has been predicted. I2C genes may be inserted in a cosmid  
 CC vector for expression in a tomato transgenic plant, to confer  
 CC disease-resistance, or may be used as restriction fragment length  
 CC polymorphism probes for screening for and selective breeding of  
 CC tomato or Solanaceae plants with disease-resistance.  
 XX  
 SQ Sequence 1220 AA;  
 SQ  
 Query Match 33.3%; Score 8; DB 17; Length 1220;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 17 EQLRTLPL 24  
 |||||  
 Db 554 eqrltlpl 561  
 |||||  
 RESULT 4  
 ID AAW03665 standard; Protein: 1240 AA.  
 XX  
 XX AAW03665;  
 AC  
 XX  
 DT 22-FEB-1997 (first entry)  
 XX  
 XX I2C-2 protein conferring Fusarium wilt disease-resistance.  
 DE  
 XX  
 KW Tomato; I2C-2; Fusarium wilt; disease-resistance; Fusarium oxysporum;  
 KW P-loop; leucine-rich repeat; transgenic plant; screening;  
 KW restriction fragment length polymorphism; crop improvement;  
 KW Solanaceae.  
 XX  
 OS Lycopersicon esculentum.  
 XX

FH Key Location/Qualifiers  
 FT Domain 201..208  
 FT /note= "P-loop"  
 FT 273..282  
 FT Domain  
 FT /note= "Conserved motif of unknown function"  
 FT 304..311  
 FT Domain  
 FT /note= "Conserved motif of unknown function"  
 FT 367..372  
 FT Domain  
 FT /note= "Conserved motif of unknown function"  
 FT 387..388  
 FT Misc-difference  
 FT /note= "Conserved motif of unknown function"  
 FT 410..415  
 FT Domain  
 FT /note= "Conserved motif of unknown function"  
 FT 488..497  
 FT Domain  
 FT /note= "Conserved motif of unknown function"  
 XX  
 XX W09632007-A1.  
 XX  
 XX 17-OCT-1996.  
 PD  
 XX  
 PF 15-APR-1996; 96WO-US05272.  
 XX  
 PR 13-APR-1995; 95IL-0113373.  
 XX  
 XX (YEDA ) YEDA RES & DEV CO LTD.  
 PA (YISS ) YISSUM RES & DEV CO.  
 PA  
 PI Eshed Y, Fluhr R, Ori N, Paran I, Zamir D;  
 DR WPI: 1996-476739/47.  
 DR N-PSDB; AAT42135.  
 XX  
 PT New DNA conferring resistance to Fusarium sp. - used for producing  
 PT disease-resistant tomato plants and for identifying new resistance  
 PT genes and diseases  
 PS  
 PS Claim 1; Fig 4A-C; 50pp; English.  
 XX  
 CC This I2C-2 protein is encoded by a sequence from the I2C multigene  
 CC family from the I2 Fusarium wilt resistance locus of tomato, and  
 CC confers resistance to Fusarium oxysporum f.sp. Lycopersici race-2.  
 CC This sequence and I2C-1 (AAW03664) are encoded by genes from a locus  
 CC completely linked to I2, and show structural similarity with other  
 CC resistance proteins. The protein has a conserved N-terminal  
 CC nucleotide-binding domain (the P-loop) and 5 other conserved  
 CC domains of unknown function. At least half the C-terminus is  
 CC composed of leucine-rich repeats, which may be responsible for  
 CC specificity of interaction, either with a pathogen protein  
 CC component, or with downstream factors involved with signal  
 CC transduction. There does not appear to be a transmembrane domain,  
 CC indicating an intracellular location. I2C genes may be inserted  
 CC in a cosmid vector for expression in a tomato transgenic plant, to  
 CC confer disease-resistance, or may be used as restriction fragment  
 CC length polymorphism probes for screening for and selective breeding  
 CC of tomato or Solanaceae plants with disease-resistance.  
 XX  
 SQ Sequence 1240 AA;  
 SQ  
 Query Match 33.3%; Score 8; DB 17; Length 1240;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 17 EQLRTLPL 24  
 |||||  
 Db 545 eqrltlpl 552  
 |||||  
 RESULT 5  
 ID AAW25157 standard; Protein: 1266 AA.  
 XX  
 XX AAW25157;  
 AC

```

XX 05-DEC-1997 (first entry)
DT Tomato immunity 2 (fungal resistance) gene product.
XX
XX Tomato; transgenic plant; disease resistance; wilt inducing fungi;
KW amplified fragment length polymorphism; AFLP; Immunity 2; I-2 gene;
KW Fusarium oxysporum; genetic engineering.
XX
XX Lycopersicon esculentum.
OS
XX W09706259-A2.
PN
XX 20-FEB-1997.
PD
XX 06-AUG-1996; 96WO-EP03480.
PE
XX 07-AUG-1995; 95EP-0401849.
PR
XX (KEYG-) KEYGENE NV.
PA
XX Simons G, Vos P, Zabeau M;
PI WPI; 1997-154265/14.
DR N-PSDB; AAT79882.
XX
XX New immunity-2 resistance gene, imparting pathogen resistance to
PT plants - used to transform plants, esp. to protect against
PT wilt-inducing fungi
XX
XX Claim 41; Fig 6a-d; 61pp; English.
PS
XX AAW25157 shows the product of the tomato immunity 2 (I-2) gene. This
CC gene was used to produce transgenic plants that are resistant to
CC wilt-inducing fungi, e.g. Fusarium 2, especially F. oxysporum f.sp.
CC Lycopersici race 2. Plants that may be transformed include tomato,
CC melon, tobacco, Arabidopsis, aubergine and potato.
CC
XX Sequence 1266 AA;
SQ

Query Match 33.3%; Score 8; DB 18; Length 1266;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 EOLRTLLP 24
DB 547 eqLrlllp 554

```

```

XX (KEYG-) KEYGENE NV.
PA
XX Haring MA, Cornelissen BJC, Mes JJ, Simons AFM;
PI WPI; 2000-516034/47.
DR N-PSDB; AAA59332.
XX
XX New I-2 resistance gene tissue-specific regulatory sequence useful in
PT plant resistance mechanisms against plant pathogens such as fungi
PT
XX Disclosure; Page 26-31; 47pp; English.
XX
XX The present sequence represents I-2 resistance protein. The specification
CC describes nucleotide sequences which have a regulatory activity on the
CC transcription of the I-2 resistance gene in plant host cells.
CC The transgenic plants, especially tomato, egg plant, potato, melon,
CC tobacco and Arabidopsis, are capable of expressing a gene mediating
CC resistance to a plant pathogen, such as fungi, in a tissue-specific
CC manner. The plant is capable of preventing infection by a plant
CC pathogen, such as fungi. Inserting the regulatory activity polynucleotide
CC into plant cell genomes is useful for providing plants with reduced
CC susceptibility to plant pathogens, especially for protecting plants
CC in cultivation.
CC
XX Sequence 1266 AA;
SQ

Query Match 33.3%; Score 8; DB 21; Length 1266;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 EOLRTLLP 24
DB 547 eqLrlllp 554

```

```

RESULT 6
AAB07754
ID AAB07754 standard; Protein; 1266 AA.
XX
AC AAB07754;
XX
DT 07-NOV-2000 (first entry)
XX
DE Amino acid sequence of the protein encoded by the I-2 resistance gene.
XX
KW Regulatory activity; transcription; I-2 resistance gene; tomato;
KW egg plant; potato; melon; tobacco; Arabidopsis; plant pathogen; fungi;
KW tissue-specific.
XX
OS Fusarium oxysporum.
XX
FN EPI024196-A1.
XX
PD 02-AUG-2000.
XX
PE 29-JAN-1999; 99EP-0400212.
XX
PR 29-JAN-1999; 99EP-0400212.

```

```

RESULT 7
AAG43791
ID AAG43791 standard; Protein; 135 AA.
XX
AC AAG43791;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54775.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.

```

PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 27-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136392.  
PR 01-JUN-1999; 99US-0136782.  
PR 03-JUN-1999; 99US-0137222.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137502.  
PR 08-JUN-1999; 99US-0137724.  
PR 10-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139451.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159325.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.

PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 29.2%; Score 7; DB 21; Length 135;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSKERLR 8  
|||||  
DB 129 sskerklr 135

RESULT 8  
AAG43790  
ID AAG43790 standard; Protein; 250 AA.

AC AAG43790;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 54774.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159638.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161040.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match 29.2%; Score 7; DB 21; Length 250;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 SSKKKR 8  
Db 244 sskkkir 250  
  
RESULT 9  
ID AAG15758 standard; Protein; 291 AA.  
XX AAG15758;  
AC AAG15758;  
DT 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 16135.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-01334256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154179.  
PR 22-SEP-1999; 99US-0154179.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 29.2%; Score 7; DB 21; Length 291;  
Best Local Similarity 100.0%; Pred. No. 9.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LRRERIK 13  
DB 75 Lrrerik 81

## RESULT 10

AAAG15757  
ID AAAG15757 standard; Protein; 341 AA.

XX AC AAAG15757;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 16134.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

XX OS Arabidopsis thaliana.

XX OS EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139819.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 15-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161923.  
PR 29-OCT-1999; 99US-0162142.

Query Match 29.2%; Score 7; DB 21; Length 341;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 LRRERIK 13  
Db 125 LRRERIK 131  
RESULT 11  
AAG15756  
ID AAG15756 standard; Protein: 443 AA.  
XX  
AC AAG15756;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 16133.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128745.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130440.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134376.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.



PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144003.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145919.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 06-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0148368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149923.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 29.2%; Score 7; DB 21; Length 443;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 7 LRRERIK 13  
|||||  
Db 227 Lrrerik 233

```

RESULT 12
ID AAM02605 standard: Protein; 150 AA.
XX AC AAM02605;
XX DT 23-OCT-1996 (first entry)
XX DE Human leukotriene C4 synthase.
XX DE
XX DE Human; leukotriene C4; synthase, conjugation; reduced glutathione;
XX DE 5,6-oxido-7,9-E-11,14-Z-eicosatetraenoic acid;
XX DE 5,6-hydroxy-6R-glutathionyl-7,9-E-11,14-Z-eicosatetraenoic acid;
XX DE myelocytic KG-1 cell; antisense; host cell; expression vector;
XX DE modulation; synthesis; treatment; inflammation; cardiac ischemia;
XX DE anaphylactic shock; cold; exercise; aspirin; induced asthma;
XX DE allergic rhinitis; antibody; detection; aberrant tissue;
XX DE malignancy; assay; modulator.
XX OS Homo sapiens.
XX OS
XX FH Key Location/Qualifiers
XX FH Domain 25..58
XX FT /note="extramembrane loop domain I"
XX FT Domain 90..113
XX FT /note="extramembrane loop domain II"
XX FT Domain 136..150
XX FT /note="extramembrane loop domain IIT"
XX PN W09532280-A1.
XX PD 30-NOV-1995.
XX XX
XX PF 16-MAY-1995; 95WO-US06173.
XX PR 20-MAY-1994; 94US-0246991.
XX PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX PI Austen KF, Lam BK, Penrose JF;
XX DR WPI: 1996-020576/02.
XX DR N-PSDB: AAT132812.
XX PT Human leukotriene C4 synthase DNA - useful for diagnosis and
XX PT therapy, partic. for inflammatory conditions
XX PS Claim 3; Fig 2; 65pp; English.
XX XX
XX CC The present sequence is the human leukotriene C4 (LTC4)
XX CC synthase, which is capable of conjugating reduced glutathione with
XX CC 5,6-oxido-7,9-E-11,14-Z-eicosatetraenoic acid to form
XX CC 5S-hydroxy-6R-glutathionyl-7,9-E-11,14-Z-eicosatetraenoic acid.
XX CC The LTC4 synthase DNA was obtd. from myelocytic KG-1 cell RNA.
XX CC A sequence antisense to the, or a host cell transformed with an
XX CC expression vector contg., the synthase cDNA can be used to
XX CC modulate the synthesis of LTC4 in the treatment of patients
XX CC with inflammatory conditions, e.g. cardiac ischaemia, anaphylactic
XX CC shock, (cold, exercise or aspirin)-induced asthma or allergic
XX CC rhinitis. Anti-synthase antibodies can be used to detect LTC4
XX CC synthase in aberrant tissues such as malignancies, while the
XX CC recombinant cells may also be used in assays to detect modulators
XX CC of LTC4 synthase.
XX XX
XX SQ Sequence 150 AA;

```

```

Query Match 25.0%; Score 6; DB 17; Length 150;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 19 LRTLTP 24
    |||||
DB 143 lrtllp 148

```

```

RESULT 13
ID AAR90285 standard: Protein; 150 AA.
XX AC AAR90285;
XX DT 02-JUL-1996 (first entry)
XX DE Human leukotriene C4 synthase.
XX DE
XX DE Leukotriene C4 synthase; antisense oligonucleotide; chronic asthma;
XX DE treatment.
XX OS Homo sapiens.
XX OS
XX PN W09533839-A1.
XX PD 14-DEC-1995.
XX XX
XX PF 19-MAY-1995; 95WO-US06137.
XX PR 03-JUN-1994; 94US-0254354.
XX PA (SEAR ) SEARLE & CO G D.
XX PI Creely DP, Hauser SD, Welsch DJ;
XX DR WPI: 1996-040240/04.
XX DR N-PSDB: AAT11631.
XX PT DNA encoding human leukotriene C4 synthase - and anti-sense
XX PT oligo-nucleotide(s), useful to treat leukotriene C4
XX PT synthase-mediated conditions, e.g. chronic asthma
XX PS Disclosure: Page 50; 70pp; English.
XX XX
XX CC The nucleic acid encoding the human leukotriene C4 synthase can be
XX CC used to produce the leukotriene C4 synthase by standard recombinant
XX CC techniques. Non-ionic antisense oligonucleotides may be used to
XX CC prevent the production of leukotriene C4 synthase polypeptides and
XX CC are useful in the treatment of leukotriene C4 and leukotriene D4
XX CC mediated conditions such as chronic asthma. The antisense
XX CC oligonucleotides have increased resistance to nuclease digestion and
XX CC can cross the cell membrane of a viable cell.
XX XX
XX SQ Sequence 150 AA;

```

```

Query Match 25.0%; Score 6; DB 17; Length 150;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 LRTLTP 24
    |||||
DB 143 lrtllp 148

```

```

RESULT 14
ID AAG54945 standard: Protein; 152 AA.
XX AC AAG54945;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SRQ ID NO: 70291.
XX KW Protein identification; signal transduction pathway; metabolic pathway;

```

KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147360.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 25.0%; Score 6; DB 21; Length 152;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SKEKLR 8  
|||||  
DB 98 skeklr 103

RESULT 15  
AAG32041  
ID AAG32041 standard; Protein; 220 AA.  
XX

AC AAG32041;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 38581.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

OS Arabidopsis thaliana.

XX EPI033405-A2.  
XX

PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.

```
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-015832.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 25.0%; Score 6; DB 21; Length 220;
Best Local Similarity 100.0%; Pred.No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 EQURTL 22
Db 102 eqIRtl 107

Search completed: September 15, 2001, 12:47:12
Job time: 162 sec
```

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

CM protein - protein search, using SW model

Run on: September 15, 2001, 12:47:52 ; Search time 35.36 Seconds  
(without alignments)  
13.975 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_163  
Perfect score: 24  
Sequence: 1 HSSKEKLRRRIKVCCEQLFTLLP 24

Scoring table:  
OLIGO  
Gapop 60.0 , Capext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents, AA:\*  
1: /cgnl\_7/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgnl\_7/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgnl\_7/prodata/1/iaa/5A\_COMB.pep:\*  
4: /cgnl\_7/prodata/1/iaa/5B\_COMB.pep:\*  
5: /cgnl\_7/prodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgnl\_7/prodata/1/iaa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	33.3	1220	3 US-08-930-996A-2	Sequence 2, Appli
2	8	33.3	1240	3 US-08-930-996A-4	Sequence 4, Appli
3	7	29.2	52	3 US-08-630-916A-76	Sequence 76, Appli
4	6	25.0	13	4 US-08-986-837-8	Sequence 8, Appli
5	6	25.0	79	2 US-08-611-510-6	Sequence 6, Appli
6	6	25.0	147	4 US-08-986-837-2	Sequence 2, Appli
7	6	25.0	150	2 US-08-254-354-2	Sequence 2, Appli
8	6	25.0	150	5 PCT-US95-06137-2	Sequence 11, Appli
9	6	25.0	268	3 US-08-871-483-11	Sequence 1, Appli
10	6	25.0	866	1 US-08-100-692-1	Sequence 1, Appli
11	6	25.0	866	2 US-08-674-030-1	Sequence 8, Appli
12	6	25.0	874	3 US-08-247-904B-8	Sequence 21, Appli
13	6	25.0	874	2 US-08-767-942A-21	Sequence 17, Appli
14	5	20.8	6	1 US-07-944-143C-17	Sequence 17, Appli
15	5	20.8	6	5 PCT-US93-08214-17	Sequence 36, Appli
16	5	20.8	6	5 PCT-US93-12679-5	Sequence 36, Appli
17	5	20.8	7	2 US-08-105-416-36	Sequence 36, Appli
18	5	20.8	7	2 US-08-473-656A-36	Sequence 14, Appli
19	5	20.8	10	1 US-08-222-851-14	Sequence 9, Appli
20	5	20.8	10	1 US-08-440-504A-14	Sequence 14, Appli
21	5	20.8	10	1 US-08-440-504A-14	Sequence 13, Appli
22	5	20.8	10	5 PCT-US94-12985-13	Sequence 9, Appli
23	5	20.8	14	4 US-08-433-613-9	Sequence 48, Appli
24	5	20.8	14	4 US-08-433-613-9	Sequence 16, Appli
25	5	20.8	16	1 US-07-944-143C-16	Sequence 16, Appli
26	5	20.8	16	5 PCT-US93-08214-16	Sequence 1, Appli
27	5	20.8	17	1 US-07-976-872B-1	Sequence 1, Appli

28	5	20.8	17	1	US-07-976-872B-2	Sequence 2, Appli
29	5	20.8	17	1	US-07-976-872B-3	Sequence 3, Appli
30	5	20.8	17	1	US-07-976-872B-4	Sequence 4, Appli
31	5	20.8	17	1	US-08-105-416-12	Sequence 12, Appli
32	5	20.8	17	1	US-08-105-416-13	Sequence 13, Appli
33	5	20.8	17	1	US-08-105-416-18	Sequence 18, Appli
34	5	20.8	17	1	US-08-105-416-19	Sequence 19, Appli
35	5	20.8	17	1	US-08-105-416-20	Sequence 20, Appli
36	5	20.8	17	1	US-08-105-416-27	Sequence 27, Appli
37	5	20.8	17	1	US-08-105-416-29	Sequence 29, Appli
38	5	20.8	17	1	US-08-105-416-30	Sequence 30, Appli
39	5	20.8	17	1	US-08-105-416-37	Sequence 37, Appli
40	5	20.8	17	1	US-08-105-416-38	Sequence 38, Appli
41	5	20.8	17	2	US-08-473-656A-12	Sequence 12, Appli
42	5	20.8	17	2	US-08-473-656A-13	Sequence 13, Appli
43	5	20.8	17	2	US-08-473-656A-18	Sequence 18, Appli
44	5	20.8	17	2	US-08-473-656A-19	Sequence 19, Appli
45	5	20.8	17	2	US-08-473-656A-20	Sequence 20, Appli

## ALIGNMENTS

```

RESULT 1
US-08-930-996A-2
; Sequence 2, Application US/08930996A
; Patent No. 6100449
;
; GENERAL INFORMATION:
; APPLICANT: FLOHR, Robert
; APPLICANT: ESHED, Yuval
; APPLICANT: ORI, Naomi
; APPLICANT: PARAN, Ilan
; APPLICANT: ZAMIR, Daniel
; TITLE OF INVENTION: A GENE FAMILY FROM THE 12 FUSARIUM RESISTANCE
; LOCUS OF TOMATO AND USE THEREOF FOR TRANSFORMATION AND
; TITLE OF INVENTION: SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROWDY AND NEWMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/930,996A
; FILING DATE: 09-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05272
; FILING DATE: 15-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 113,373
; FILING DATE: 13-APR-1995
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1220 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-930-996A-2

```

Query Match 33.3%; Score 8; DB 3; Length 1220;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 554 EQLRTLIP 561

RESULT 2

US-08-930-996A-4

Sequence 4, Application US/08930996A

Patent No. 6100449

GENERAL INFORMATION:

APPLICANT: FLUHR, Robert

APPLICANT: ESHED, Yuval

APPLICANT: ORI, Naomi

APPLICANT: PARAN, Daniel

TITLE OF INVENTION: A GENE FAMILY FROM THE 12 PUSARIUM RESISTANCE

TITLE OF INVENTION: LOCUS OF TOMATO AND USE THEREOF FOR TRANSFORMATION AND

NUMBER OF SEQUENCES: SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEWMARK

STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/930,996A

FILING DATE: 09-DEC-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/05272

FILING DATE: 15-APR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: IL 113,373

FILING DATE: 13-APR-1995

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1240 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-930-996A-4

Query Match 33.3%; Score 8; DB 3; Length 1240;

Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 EQLRTLIP 24

Db 545 EQLRTLIP 552

RESULT 3

US-08-630-916A-76

Sequence 76, Application US/08630916A

Patent No. 6011137

GENERAL INFORMATION:

APPLICANT: Pitozzi, Gregorio

APPLICANT: Kay, Brian K.

APPLICANT: Fowlkes, Dana M.

TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL

TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME

NUMBER OF SEQUENCES: 124

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennile & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630,916A

FILING DATE: 03-APR-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MISROCK, S. LESLIE

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-203

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 896-8864/9741

INFORMATION FOR SEQ ID NO: 76:

SEQUENCE CHARACTERISTICS:

LENGTH: 52 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-630-916A-76

Query Match 29.2%; Score 7; DB 3; Length 52;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKEXLR 8

Db 30 SSKEXLR 36

RESULT 4

US-08-986-837-8

Sequence 8, Application US/08986837

Patent No. 6221676

GENERAL INFORMATION:

APPLICANT: Lam, Bing K.

APPLICANT: Pentose, John F.

TITLE OF INVENTION: DNA ENCODING HUMAN LEUKOTRIENE C4 SYNTHASE,

TITLE OF INVENTION: POLYPEPTIDES, AND USES THEREFOR

FILE REFERENCE: Docket No. 6221676 0092662-0019 (BWH269)

CURRENT APPLICATION NUMBER: US/08/986,837

CURRENT FILING DATE: 1997-12-18

EARLIER APPLICATION NUMBER: 08/246,991

EARLIER FILING DATE: 1994-05-20

NUMBER OF SEQ ID NOS: 9

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 8

LENGTH: 15

TYPE: PRT

ORGANISM: Homo sapiens

US-08-986-837-8

Query Match 25.0%; Score 6; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 3.5;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLIP 24

Db 8 LRTLIP 13

RESULT 5

US-08-611-510-6

Sequence 6, Application US/08611510

Patent No. 5824529



GENERAL INFORMATION:  
APPLICANT: Chang, Zhiyuh  
TITLE OF INVENTION: Method For Cloning And  
TITLE OF INVENTION: Producing The Pshai Restriction  
TITLE OF INVENTION: Endonuclease  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Gregory D. Williams; NEW ENGLAND  
ADDRESSEE: BIOLABS, INC.  
STREET: 32 Tozer Road  
CITY: Beverly  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 01915  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version  
SOFTWARE: #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/611,510  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: WILLIAMS, GREGORY D.  
REGISTRATION NUMBER: 30901  
REFERENCE/DOCKET NUMBER: NEB-119  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508) 927-1705  
TELEFAX: (508) 927-1705  
INFORMATION FOR SEQ. ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 79 amino acids  
TYPE: amino acid  
STRANDEDNESS: not applicable  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-611-510-6

Query Match 25.0%; Score 6; DB 2; Length 79;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERI 12  
12 LRRERI 12  
DB 16 LRRERI 21

RESULT 6  
US-08-986-837-2  
Sequence 2, Application US/08986837  
GENERAL INFORMATION:  
APPLICANT: Lam, Bing K.  
APPLICANT: Penrose, John F.  
APPLICANT: Frank, Austen K.  
TITLE OF INVENTION: DNA ENCODING HUMAN LEUKOTRIENE C4 SYNTHASE,  
TITLE OF INVENTION: POLYPEPTIDES, AND USES THEREFOR  
FILE REFERENCE: Docket No. 6221676 0092662-0019 (BWH269)  
CURRENT APPLICATION NUMBER: US/08/986,837  
CURRENT FILING DATE: 1997-12-18  
EARLIER FILING DATE: 1994-05-20  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2  
LENGTH: 147  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-986-837-2

Query Match 25.0%; Score 6; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLLP 24  
19 LRTLLP 24  
DB 143 LRTLLP 148

RESULT 7  
US-08-254-354-2  
Sequence 2, Application US/08254354  
Patent No. 5952210  
GENERAL INFORMATION:  
APPLICANT: Greely, David P.  
APPLICANT: Hauser, Scott D.  
APPLICANT: Welsch, Dean J.  
TITLE OF INVENTION: Nucleic Acids and Expression Vectors  
TITLE OF INVENTION: Encoding Human Leukotriene C4 Syn-  
TITLE OF INVENTION: Oligonucleotides, and Methods of Use  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scott B. Feder, G. D. Searle & Co., Corporate  
ADDRESSEE: Patent Dept.  
STREET: P. O. Box 5110  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60680  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/254,354  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Feder, Scott B.  
REGISTRATION NUMBER: 33,129  
REFERENCE/DOCKET NUMBER: 2800  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (708)470-6501  
TELEFAX: (708)470-6881  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 150 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-254-354-2

Query Match 25.0%; Score 6; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLLP 24  
19 LRTLLP 24  
DB 143 LRTLLP 148

RESULT 8  
PCT-US95-06137-2  
Sequence 2, Application PC/TUS9506137  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Nucleic Acids and Expression Vectors Encoding  
TITLE OF INVENTION: Human leukotriene C4 Synthase, Antisense Oligonucleotides,  
TITLE OF INVENTION: and Methods of Use

NUMBER OF SEQUENCES: 12  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/06137  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/254354  
FILING DATE: 03-JUN-1994  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 150 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-06137-2

Query Match 25.0%; Score 6; DB 5; Length 150;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLLP 24  
Db 143 LRTLLP 148

RESULT 9  
US-08-871-483-11  
Sequence 11, Application US/08871483  
Patent No. 6083715  
GENERAL INFORMATION:  
APPLICANT: Georgiou, George  
APPLICANT: Qui, Ji  
APPLICANT: Bessette, Paul  
TITLE OF INVENTION: METHODS FOR PRODUCING HETEROLOGOUS  
TITLE OF INVENTION: DISULFIDE BOND-CONTAINING PEPTIDES IN BACTERIAL CELLS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/871,483  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitchell, Barbara S.  
REGISTRATION NUMBER: 33,928  
REFERENCE/DOCKET NUMBER: UTSB:614  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-871-483-11

Query Match 25.0%; Score 6; DB 3; Length 268;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0,  
OY 18 QLRTLL 23  
Db 166 QLRTLL 171

RESULT 10  
US-08-100-692-1  
Sequence 1, Application US/08100692  
Patent No. 5532348  
GENERAL INFORMATION:  
APPLICANT: Hulbregtse, Jon M.  
APPLICANT: Scheffner, Martin  
TITLE OF INVENTION: E6 ASSOCIATED PROTEIN AND METHODS OF USE  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza, Stewart Street Tower  
CITY: San Francisco  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94105-1492  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/100,692  
FILING DATE: 19930730  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 15280-91  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 866 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-100-692-1

Query Match 25.0%; Score 6; DB 1; Length  
Best Local Similarity 100.0%; Pred. No. 11e+02;  
Matches 6; Conservative 0; Mismatches 0;

OY 2 SSKEKL 7  
Db 844 SSKEKL 849

RESULT 11  
US-08-674-030-1  
Sequence 1, Application US/0867403  
Patent No. 5914389  
GENERAL INFORMATION:  
APPLICANT: Hulbregtse, Jon M.  
APPLICANT: Scheffner, Martin  
TITLE OF INVENTION: E6 ASSOCIATED PROTEIN AND METHODS OF USE  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza, Stewart Street Tower  
CITY: San Francisco  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94105-1492  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/674,030  
FILING DATE: 19930730  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 15280-91  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 467-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 866 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-674-030-1

TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend Hourie and Crew  
STREET: One Market Plaza, Steuart Street Tower  
CITY: San Francisco  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94105-1492

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/674,030  
FILING DATE: 01-JUL-1996  
CLASSIFICATION: 530

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/100,692  
FILING DATE: 30-JUL-1993

## ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 15280-91

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 467-9600  
TELEFAX: (415) 543-5043

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 866 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-674-030-1

Query Match 25.0%; Score 6; DB 2; Length 866;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKEXL 7  
|||||

DB 844 SSKEXL 849

## RESULT 12

US-08-247-904B-8

Sequence 8, Application US/08247904B

Patent No. 5981699

## GENERAL INFORMATION:

APPLICANT: Rolfe, Mark

APPLICANT: Eckstein, Jens W.

APPLICANT: Draetta, Giulio

TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme

NUMBER OF SEQUENCES: 17

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley, Hoag & Elliot

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/247,904B

FILING DATE: 23-MAY-1994

CLASSIFICATION: 530

## ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew

REGISTRATION NUMBER: 36,777

REFERENCE/DOCKET NUMBER: 1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000

TELEFAX: (617) 832-7000

INFORMATION FOR SEQ ID NO: 8:

## SEQUENCE CHARACTERISTICS:

LENGTH: 874 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-247-904B-8

Query Match 25.0%; Score 6; DB  
Best Local Similarity 100.0%; Pred. No. 1.1  
Matches 6; Conservative 0; Mismatches

QY 2 SSKEXL 7  
|||||

DB 852 SSKEXL 857

## RESULT 13

US-08-767-942A-21

Sequence 21, Application US/08767942A

Patent No. 6068982

## GENERAL INFORMATION:

APPLICANT: Rolfe, Mark

APPLICANT: Chiu, M. Isabel

APPLICANT: Berlin, Vivian

APPLICANT: Damagnez, Veronique

APPLICANT: Draetta, Giulio

APPLICANT: Guillaume, Cottarel

TITLE OF INVENTION: UBUIQUITIN CONJUGATING ENZYMES

NUMBER OF SEQUENCES: 45

## CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/767,942A

FILING DATE: 17-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: MIV-029.04

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-832-7000

TELEFAX: 617-832-1000

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 874 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-767-942A-21

Query Match 25.0%; Score 6; DB 3; Length 874;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0,

QY 2 SSKKL 7  
DB 852 SSKKL 857

RESULT 14  
US-07-944-143C-17  
Sequence 17, Application US/07944143C  
Patent No. 5719064  
GENERAL INFORMATION:  
APPLICANT: Scofield, R. Hal  
APPLICANT: Harley, John B.  
TITLE OF INVENTION: Peptide Diagnostics and Therapeutics for  
Spondyloarthropathies  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/944,143C  
FILING DATE:  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMR138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)873-8794  
TELEFAX: (404)873-8795  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Salmonella typhimurium  
US-07-944-143C-17

Query Match 20.8%; Score 5; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
DB 2 LRTLL 6

RESULT 15  
PCT-US93-08214-17  
Sequence 17, Application PC/TUS9308214  
GENERAL INFORMATION:  
APPLICANT: Oriana Medical Research Foundation  
TITLE OF INVENTION: Peptide Diagnostics and Therapeutics  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET 1100 Peachtree Street, Suite 2800

CITY: Atlanta  
STATE: Georgia  
COUNTRY: United States  
ZIP: 30309-4530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08214  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMR138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 815-6555  
TELEFAX: (404) 815-6555  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Salmonella typhimurium  
PCT-US93-08214-17

Query Match 20.8%; Score 5; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 19 LRTLL 23  
DB 2 LRTLL 6

Search completed: September 15, 2001, 12:47:53  
Job time: 183 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:48:52 : Search time 45.39 Seconds  
(without alignments)  
40.277 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_163  
Perfect score: 24  
Sequence: 1 HSSKELRRERIKYCCQLRLTLP 24

Scoring table:  
Gapop 60.0, Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size: 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	33.3	1220	2	T06403 resistance complex
2	8	33.3	1240	2	T06404 resistance complex
3	7	29.2	250	2	H85067 hypothetical prote
4	7	29.2	327	2	E64188 dipeptide transpor
5	7	29.2	443	2	T48593 hypothetical prote
6	7	29.2	486	2	F64204 spermatine/putresc
7	7	29.2	748	2	T30634 hypothetical prote
8	7	29.2	770	1	A44337 kinesin-related pr
9	7	29.2	791	2	T20815 hypothetical prote
10	7	29.2	1341	2	S66835 probable membrane
11	6	25.0	68	1	IABY3 proteinase A inhib
12	6	25.0	81	2	S61847 hrpx protein - pse
13	6	25.0	87	2	C82334 probable host fact
14	6	25.0	90	2	H69159 hypothetical prote
15	6	25.0	91	2	D64066 host factor I - Ha
16	6	25.0	93	2	T40184 conserved hypothet
17	6	25.0	94	2	T12719 hypothetical prote
18	6	25.0	109	2	F82067 probable anti-sim
19	6	25.0	135	2	T19002 hypothetical prote
20	6	25.0	145	2	C71378 hypothetical prote
21	6	25.0	150	2	I38595 leucotoiene-C4 syn
22	6	25.0	159	2	G84974 transcription elon
23	6	25.0	160	2	T24339 hypothetical prote
24	6	25.0	160	2	S52889 probable membrane
25	6	25.0	177	2	E71138 probable phosphori
26	6	25.0	181	1	JVBPNL DNA-packaging prot
27	6	25.0	181	2	F64788 DNA packaging prot
28	6	25.0	181	2	G85689 hypothetical prote
29	6	25.0	184	2	T38315 very hypothetical

30	6	25.0	189	2	G64909 DNA packaging prot
31	6	25.0	211	2	F75474 hypothetical prote
32	6	25.0	215	2	E83564 hypothetical prote
33	6	25.0	218	2	C73594 ATP phosphoribosyl
34	6	25.0	219	2	F71155 hypothetical prote
35	6	25.0	222	2	S71231 geranylgeranyl pyr
36	6	25.0	222	2	C72400 conserved hypothet
37	6	25.0	232	2	C75261 probable iron depe
38	6	25.0	239	2	A84091 hypothetical prote
39	6	25.0	242	2	A75023 hypothetical prote
40	6	25.0	242	2	S39642 motility protein h
41	6	25.0	249	2	B75110 lipote-protein ii
42	6	25.0	250	2	T44227 hypothetical prote
43	6	25.0	250	2	T44041 glycoprotein L [lm
44	6	25.0	250	2	J02165 glycoprotein gl pr
45	6	25.0	254	2	D64560 outer membrane pro

## ALIGNMENTS

```

RESULT 1
T06403
Resistance complex protein I2C-1 - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 17-Nov-2000
C:Accession: T06403
R:Ori, N: Eshed, Y.: Paran, I.: Prestling, G.: Aviv, D.: Tanksley, S.: Zamir, D.: Flu
Plant Cell 9, 521-532, 1997
A>Title: The I2C family from the wilt disease resistance locus I2 belongs to the nucl
A:Reference number: Z15652; MUID:97290204
A:Accession: T06403
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1220 <ORI>
A:Cross-references: EMBL:AF004878; NID:g2258314; PIDN:AAB63274.1; PID:g2258315
C:Genetics:
A:Gene: I2C-1
A:Map position: 11
C:Function:
A>Description: confers resistance against Fusarium oxysporum
C:Superfamily: disease resistance protein RPS2; leucine-rich alpha-2-glycoprotein rep

Query Match      33.3%; Score 8; DB 2; Length 1220;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 EQLRLTLP 24
DB 554 EQLRLTLP 561

RESULT 2
T06404
Resistance complex protein I2C-2 - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 17-Nov-2000
C:Accession: T06404
R:Ori, N: Eshed, Y.: Paran, I.: Prestling, G.: Aviv, D.: Tanksley, S.: Zamir, D.: Flu
Plant Cell 9, 521-532, 1997
A>Title: The I2C family from the wilt disease resistance locus I2 belongs to the nucl
A:Reference number: Z15652; MUID:97290204
A:Accession: T06404
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1240 <ORI>
A:Cross-references: EMBL:AF004879; NID:g2258316; PIDN:AAB63275.1; PID:g2258317
C:Genetics:
A:Gene: I2C-2
A:Map position: 11
C:Function:
A>Description: confers resistance against Fusarium oxysporum

```

C:Superfamily: disease resistance protein RPS2; leucine-rich alpha-2-glycoprotein repeat

Query Match 33.3%; Score 8; DB 2; Length 1240;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 EQLRTLLP 24  
|||||  
DB 545 EQLRTLLP 552

RESULT 3

H85067  
hypothetical protein AT4G05400 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001  
C:Accession: H85067  
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488  
A:Accession: H85067  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-250 <STO>  
A:Cross-references: GB:NC\_001268; NID:g7267300; PIDN:CA881082.1; GSPDB:GN00140  
C:Genetics:  
A:Gene: AT4G05400  
A:Map position: 4  
C:Superfamily: Arabidopsis thaliana hypothetical protein FJ77.80

Query Match 29.2%; Score 7; DB 2; Length 250;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKERLR 8  
|||||  
DB 244 SSKERLR 250

RESULT 4

E64188  
dipeptide transport ATP-binding protein dppef - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 20-Aug-1999  
C:Accession: E64188  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: E64188  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-327 <TIGR>  
A:Cross-references: GB:U3798; GB:LA2023; NID:g1574110; PIDN:AMC22837.1; PID:g1574111; T  
C:Genetics:  
A:Gene: dppef  
A:Function:  
A:Description: probably responsible for energy-coupling to the transport system  
C:Superfamily: inner membrane protein malK; ATP-binding cassette homology  
C:Keywords: ATP; binding protein-dependent transport system; membrane protein; nucleotide  
P:37-237/Domain; ATP-binding cassette homology <ABC>  
E:54-62/Region; nucleotide-binding motif A (P-loop)  
F:181-185/Region; nucleotide-binding motif B

Query Match 29.2%; Score 7; DB 2; Length 327;  
Best Local Similarity 100.0%; Pred. No. 6.5;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERIK 13  
|||||  
DB 271 LRRERIK 277

RESULT 5

T48593  
hypothetical protein T22N19.110 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T48593  
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancro  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24490  
A:Accession: T48593  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-443 <BEV>  
A:Cross-references: EMBL:AL163572  
A:Experimental source: cultivar Columbia; BAC clone T22N19  
C:Genetics:  
A:Map position: 5  
A:Introns: 23/3; 196/3; 240/3  
A:Note: T22N19.110

Query Match 29.2%; Score 7; DB 2; Length 443;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERIK 13  
|||||  
DB 227 LRRERIK 233

RESULT 6

F64204  
spermidine/putrescine transport ATP-binding protein potA homolog - Mycoplasma genital  
C:Species: Mycoplasma genitalium  
C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 17-Mar-2000  
C:Accession: F64204  
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.  
M.; Fuhmann, J.; Nguyen, D.; Uitterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,  
C.A.; Venter, J.C.  
Science 270, 397-403, 1995  
A:Title: The minimal gene complement of Mycoplasma genitalium.  
A:Reference number: A64200; MUID:96026346  
A:Accession: F64204  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-486 <TIGR>  
A:Cross-references: GB:U39683; GB:LA3967; NID:g1045711; PID:g1045714; TIGR:MG042  
A:Experimental source: strain G-37  
C:Genetics:  
A:Genetic code: SGC3  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog  
C:Keywords: ATP  
F:1-351/Domain: ATP-binding cassette homology #status atypical <ABCL>

Query Match 29.2%; Score 7; DB 2; Length 486;  
Best Local Similarity 100.0%; Pred. No. 9.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KEKLRRE 10  
|||||  
DB 100 KEKLRRE 106

RESULT 7  
T30634



hypothetical protein 32L - Molluscum contagiosum virus 1  
N:Alternate names: MC032L  
C:Species: Molluscum contagiosum virus 1  
C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jul-2000  
C:Accession: T30634  
R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.  
Science 273, 813-816, 1996  
A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re  
A:Reference number: Z20876; MUID:96325459  
A:Accession: T30634  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-748 <SEN>  
A:Cross-references: EMBL:060315; NID:g1491943; PIDN:AAC55160.1; PID:g1491975  
C:Genetics:  
A:Note: MC032L  
C:Superfamily: vaccinia virus hypothetical protein 49L

Query Match 29.2%; Score 7; DB 2; Length 748;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 YCCEQLR 20  
|||||  
DB 519 YCCEQLR 525

RESULT 8  
A44337  
Kinesin-related protein KLPa - Emericella nidulans  
N:Alternate names: Kinesin-like protein, KAR3-related; KLPa protein  
C:Species: Emericella nidulans, Aspergillus nidulans  
C>Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 19-Jan-2001  
C:Accession: A44337; S24830  
R:O'Connell, M.J.; Meluh, P.B.; Rose, M.D.; Morris, N.R.  
J. Cell Biol. 120, 153-162, 1993  
A:Title: Suppression of the bimC4 mitotic spindle defect by deletion of klpA, a gene enc  
A:Reference number: A44337; MUID:93107178  
A:Accession: A44337  
A:Molecule type: mRNA  
A:Residues: 1-770 <OLC>  
A:Cross-references: GB:X64603; NID:92703; PIDN:CAA45887.1; PID:g2704  
A:Note: Sequence extracted from NCBI backbone (NCBIRP.121121)  
C:Genetics:  
A:Gene: KLPa  
A:Map position: 1  
C:Superfamily: kinesin-related protein KLPa; kinesin motor domain homology  
C:Keywords: ATP; coiled coil; microtubule binding; nucleotide binding; P-loop  
F:422-762/Domain: kinesin motor domain homology <KMOT>  
F:514-521/Region: nucleotide-binding motif A (P-loop)  
F:520/Binding site: ATP (Lys) #status Predicted

Query Match 29.2%; Score 7; DB 1; Length 770;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKLRE 10  
|||||  
DB 398 KEKLRE 404

RESULT 9  
T20815  
hypothetical protein F26D2.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 26-Oct-1999  
C:Accession: T20815; T21409  
R:McMurray, A.  
submitted to the EMBL Data Library, March 1997  
A:Reference number: Z19327  
A:Accession: T20815

A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-791 <WIL>  
A:Cross-references: EMBL:Z93377; PIDN:CAB07581.1; GSPDB:GN00023; CESP:F26D2.2  
A:Experimental source: clone F13A7  
R:McMurray, A.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19418  
A:Accession: T21409  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-791 <W12>  
A:Cross-references: EMBL:Z81513; PIDN:CAB04186.1; GSPDB:GN00023; CESP:F26D2.2  
A:Experimental source: clone F26D2  
C:Genetics:  
A:Gene: CESP:F26D2.2  
A:Map position: 5  
A:Introns: 28/3; 194/2; 195/3; 299/2; 404/1; 462/3; 500/1; 670/3; 698/3

Query Match 29.2%; Score 7; DB 2; Length 791;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 EKLRRER 11  
|||||  
DB 153 EKLRRER 159

RESULT 10  
S66835  
probable membrane protein YOL138c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein AOD1341; hypothetical protein O0483  
C:Species: Saccharomyces cerevisiae  
C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 21-Jul-2000  
C:Accession: S66835; S72030  
R:Arino, J.; Casamayor, A.; Gamo, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S66814  
A:Accession: S66835  
A:Molecule type: DNA  
A:Residues: 1-1341 <ARI>  
A:Cross-references: EMBL:Z74880; NID:g1420031; PID:e251919; PID:g1420032; MIPS:YOL138  
A:Experimental source: strain S268C  
R:Aldea, M.; Piedrafitra, L.; Casas, C.; Casamayor, A.; Khalid, H.; Balcells, L.; Arin  
Yeast 12, 1053-1058, 1996  
A:Title: Sequence analysis of a 12 801 bp fragment of the left arm of yeast chromosome  
protein and six other open reading frames.  
A:Reference number: S72030; MUID:97051593  
A:Accession: S72030  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1223, 'V', 1224-1341 <ALD>  
A:Cross-references: EMBL:X95465; NID:g1628437; PIDN:CAA64732.1; PID:g1628438  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996  
C:Genetics:  
A:Map position: 15L  
A:Note: YOL138C  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
C:Keywords: transmembrane protein  
F:254-288/Domain: WD repeat homology <WD1>  
F:365-398/Domain: WD repeat homology <WD2>  
F:1178-1194/Domain: transmembrane #status predicted <TMM>

Query Match 29.2%; Score 7; DB 2; Length 1341;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSKELR 8  
|||||  
DB 1257 SSKELR 1263

## RESULT 11

IABY3

proteinase A inhibitor 3 - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: protein IM8010.04c; protein YMR174c

C:Species: *Saccharomyces cerevisiae*

C:Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 21-Jul-2000

C:Accession: A01334; S16692; S55121

R:Biemann, K.; Montali, U.; Martin, B.; Svendsen, I.; Ottesen, M.

C:Keywords: Commun. 45, 225-235, 1980

A:Title: The amino acid sequence of proteinase A inhibitor 3 from baker's yeast.

A:Reference number: A01334

A:Accession: A01334

A:Molecule type: protein

A:Residues: 1-68 &lt;BIE&gt;

A:Note: nearly all of the inhibitory activity is present in the peptide consisting of re

R:Schu, P.; Wolf, D.H.

FEBS Lett. 283, 78-84, 1991

A:Title: The proteinase yscA-inhibitor, I(A)(3), gene. Studies of cytoplasmic proteinase

A:Reference number: S16692; MUID:91243884

A:Accession: S16692

A:Molecule type: DNA

A:Residues: 1-68 &lt;SCH&gt;

A:Cross-references: EMBL:X60050; NID:94094; PIDN:CAA42650.1; PID:94095

R:Churcher, C.M.

submitted to the EMBL Data Library, June 1995

A:Reference number: S55118

A:Accession: S55121

A:Molecule type: DNA

A:Residues: 1-68 &lt;CHU&gt;

A:Cross-references: EMBL:Z49808; NID:9854440; PIDN:CAA89907.1; PID:9854444; GSPDB:GN0001

A:Experimental source: strain AB972

C:Genetics:

A:Gene: SGD:PA13; MIPS:YMR174c

A:Cross-references: SGD:S0004786; MIPS:YMR174c

A:Map position: 13R

A:Superfamily: proteinase A inhibitor 3

C:Keywords: acetylated amino end; blocked amino end; proteinase inhibitor

F:1/Modified site: acetylated amino end (Met) #status experimental

QY 2 SSKERL 7

Db 14 SSKERL 19

## RESULT 12

hrpx protein - *Pseudomonas solanacearum*C:Species: *Pseudomonas solanacearum*

C:Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 08-Oct-1999

C:Accession: S61847

R:van Gijsegem, F.; Gough, C.; Zischek, C.; Niqueux, E.; Arlat, M.; Genin, S.; Barberis, M.O.

A:Title: The hrp gene locus of *Pseudomonas solanacearum*, which controls the production of

A:Reference number: S61846; MUID:95349395

A:Accession: S61847

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-81 &lt;VAN&gt;

A:Cross-references: EMBL:Z14056; NID:9550397; PIDN:CAA78421.1; PID:9550399

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1994

C:Genetics:

A:Gene: hrpx

Query Match 25.0%; Score 6; DB 2; Length 81;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 EQLRTL 22

Db 59 EQLRTL 64

## RESULT 13

C82334

probable host factor-I VC0347 (imported) - *Vibrio cholerae* (strain N16961 serogroup OC:Species: *Vibrio cholerae*

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: C82334

R:Heidelberg, J.F.; Elsen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers

L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A:Reference number: A82035; MUID:20406833

A:Accession: C82334

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-87 &lt;HEI&gt;

A:Cross-references: GB:AE004123; GB:AE003852; NID:9654756; PIDN:AAF93520.1; GSPDB:GN

A:Experimental source: serogroup O1, strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC0347

A:Map position: 1

C:Superfamily: host factor I

QY 7 LRRERI 12

Db 15 LRRERI 20

## RESULT 14

H69159

hypothetical protein MTH458 - *Methanobacterium thermoautotrophicum* (strain Delta H)C:Species: *Methanobacterium thermoautotrophicum*

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999

C:Accession: H69159

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T

; Qiu, D.; Spadefora, R.; Vicaire, R.; Wang, Y.; Mierzbowski, J.; Gibson, R.; Jiwani, K.I.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: fu

A:Reference number: A69000; MUID:98037514

A:Accession: H69159

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-90 &lt;MTH&gt;

A:Cross-references: GB:AE000830; GB:AE000666; NID:92621523; PIDN:AA884964.1; PID:9262

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH458

Query Match 25.0%; Score 6; DB 2; Length 90;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KEXLRR 9

Db 36 KEXLRR 41

## RESULT 15

D64066

host factor I - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 29-Sep-1999  
C:Accession: D64066  
R:Flatischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64060; MUID:95350630  
A:Accession: D64066  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-91 <TIGR>  
A:Cross-references: GB:U32724; GB:LA2023; NID:g1573378; PIDN:AAC22070.1; PID:g1573384; T  
C:Superfamily: host factor I

Query Match 25.0%; Score 6; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 LRRERI 12  
|||||  
DB 15 LRRERI 20

Search completed: September 15, 2001, 12:48:53  
Job time: 223 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:50:40 ; Search time 23.18 Seconds  
(without alignments)  
35.467 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_163  
Perfect score: 24  
Sequence: 1 HSSKEKLRERIKYCEQDLRLLP 24

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	29.2	327	1	DPEF_HAEIN
2	7	29.2	559	1	POTA_MYGE
3	7	29.2	770	1	KLPA_EMENT
4	6	25.0	68	1	IPA3_YEAST
5	6	25.0	90	1	HFO_HAEIN
6	6	25.0	93	1	YB4C_SCHPO
7	6	25.0	145	1	R17_ORYSA
8	6	25.0	145	1	Y004_TREPA
9	6	25.0	150	1	LC4S_HUMAN
10	6	25.0	159	1	GREB_BUCAT
11	6	25.0	160	1	YRTO_YEAST
12	6	25.0	177	1	PURE_PYRHO
13	6	25.0	181	1	NOHB_ECOLI
14	6	25.0	181	1	TERS_LAMB
15	6	25.0	189	1	NOHA_ECOLI
16	6	25.0	242	1	YTXE_BACSU
17	6	25.0	248	1	DSBG_ECOLI
18	6	25.0	250	1	VGIL_HSV6
19	6	25.0	250	1	VGIL_HSV6
20	6	25.0	258	1	VGIL_GPCMV
21	6	25.0	318	1	Y211_AQUAE
22	6	25.0	328	1	CEBB_CHICK
23	6	25.0	334	1	DPEF_ECOLI
24	6	25.0	346	1	HRP1_PESY
25	6	25.0	372	1	BIOF_METJA
26	6	25.0	377	1	TDP_DROME
27	6	25.0	389	1	PYRD_DROME
28	6	25.0	397	1	OSFA_DROME
29	6	25.0	415	1	CC11_YEAST
30	6	25.0	509	1	G6PD_ANASP
31	6	25.0	509	1	G6PD_NOSPU
32	6	25.0	522	1	RECN_HELPJ
33	6	25.0	524	1	RECN_HELPJ

34	6	25.0	526	1	BUTY_BOVIN
35	6	25.0	538	1	NREA_HAEIN
36	6	25.0	550	1	YMAO_YEAST
37	6	25.0	552	1	Y4HP_RHISN
38	6	25.0	556	1	PRIS_THIFE
39	6	25.0	561	1	CCB2_DROME
40	6	25.0	743	1	PMT5_YEAST
41	6	25.0	769	1	VPL_BPH6
42	6	25.0	795	1	SYTB_ECOLI
43	6	25.0	875	1	UE3A_HUMAN
44	6	25.0	885	1	UE3A_MOUSE
45	6	25.0	906	1	CAD2_HUMAN

## ALIGNMENTS

```

RESULT 1
DPEF_HAEIN STANDARD; PRT; 327 AA.
AC P45094;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DIPEPTIDE TRANSPORT ATP-BINDING PROTEIN DPEF.
GN DPEF OR H1184.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.T., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Heddlom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd."
RL Science 269:496-512(1995).
RC SEQUENCE FROM N.A.
RX STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=96134971; PubMed=8550458;
RA Preston A., Maskell D., Johnson A., Moxon E.R.;
RA "Altered lipopolysaccharide characteristic of the 169 phenotype in
Haemophilus influenzae results from mutations in a novel gene, isn."
RL J. Bacteriol. 178:396-402(1996).
CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
FOR DIPEPTIDES. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO
THE TRANSPORT SYSTEM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
(ABC TRANSPORTERS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement. (See http://www.isb.ch/announce/
or send an email to license@isb.slb.ch).
CC EMBL: U32798; AAC22837.1; -
CC DR EMBL: U17295; AAA5975.1; -
CC TIGR: H1184; -
CC InterPro: IPR001617; -

```

```

DR Pfam: PF00005; ABC_tran; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW Peptide transport; Transport; Inner membrane; ATP-binding.
FT NP_BIND 54 61
SQ SEQUENCE 327 AA; 36917 MW; 0BB0BDE197DA9BE CRC64;

Query Match 29.2%; Score 7; DB 1; Length 327;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LRRERIK 13
DB 271 LRRERIK 277

RESULT 2
POTA_MYCGE STANDARD: PRT; 559 AA.
AC P47288;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SPERMIDINE/PUTRESCINE TRANSPORT ATP-BINDING PROTEIN POTA HOMOLOG.
GN POTA OR MG042.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MYCG 33530 / G-37;
RA MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Fullmann J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RA "The minimal gene complement of Mycoplasma genitalium.";
RA Science 270:397-403(1995).
RT -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR SPERMIDINE/PUTRESCINE. PROBABLY RESPONSIBLE FOR ENERGY
CC COUPLING TO THE TRANSPORT SYSTEM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U39684; AAC71358.1; -
CC DR HSSP: P13569; INBD.
CC DR TIGR: MG042; -
CC DR InterPro: IPR001617; -
CC DR Pfam: PF00005; ABC_tran; 1.
CC DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
CC KW Transport; ATP-binding; Membrane.
CC FT NP_BIND 40 47
CC FT SEQUENCE 559 AA; 65192 MW; DB8F04B26A110F6F CRC64;

Query Match 29.2%; Score 7; DB 1; Length 559;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKLRE 10
DB 11111111

```

```

DB 173 KEKLRE 179

RESULT 3
KLPA_EMENT STANDARD: PRT; 770 AA.
ID KLPA_EMENT
AC P28739;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE KINESIN-LIKE PROTEIN KLPA.
GN KLPA.
OS Emeritella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emeritella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GB20;
RA MEDLINE=93107178; PubMed=8416986;
RA O'Connell M.J., Meluh P.B., Rose M.D., Morris N.R.;
RA "Suppression of the bimC mitotic spindle defect by deletion of Klpa,
RT a gene encoding a KAR3-related kinesin-like protein in Aspergillus
RT nidulans.";
RT J. Cell Biol. 120:153-162(1993).
RT -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X64603; CAA45867.1; -
CC DR PIR: A44337; A44337.
CC DR HSSP: P17119; 3KAR.
CC DR InterPro: IPR001752; -
CC DR Pfam: PF00225; Kinesin; 1.
CC DR PRINTS: PR00380; KINESINHEAVY.
CC DR PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
CC DR PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
CC KW Motor protein; Microtubules; ATP-binding; Coiled coil.
CC FT DOMAIN 175 425
CC FT COILED COIL (POTENTIAL).
CC FT MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
CC FT NP_BIND 514 521
CC FT SEQUENCE 770 AA; 85800 MW; EFD0F0FF39B5C7EB CRC64;

Query Match 29.2%; Score 7; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKLRE 10
DB 398 KEKLRE 404

RESULT 4
IPA3_YEAST STANDARD: PRT; 68 AA.
ID IPA3_YEAST
AC P01094;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PROTEASE A INHIBITOR 3 (PROTEINASE INHIBITOR I(A)3).
GN PAI3 OR YMR174C OR YMR8010.04C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C:
RX MEDLINE-91243884; PubMed-2037077;
RA Schu P., Wolf D.H.;
RT "The proteinase yscA-inhibitor, I13, gene. Studies of cytoplasmic
RL proteinase inhibitor deficiency on yeast physiology."
FEBS Lett. 283:78-84(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA Church C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE.
RA Biedermann K., Montali U., Martin B., Svendsen I., Ottesen M.;
RT "The amino acid sequence of proteinase A inhibitor 3 from baker's
RL yeast."
CC Carlsberg Res. Commun. 45:225-235(1980).
CC -1- FUNCTION: INHIBITOR FOR YEAST PROTEASE A (YSCA).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X60050; CAA42650.1; -
CC EMBL; 249808; CAA89907.1; -
CC PIR; A01334; IABY3.
CC PIR; S16692; S16692.
CC SGD; S0004786; PA13.
CC Protease inhibitor; Acetylation.
CC MOD_RES 1 1 ACETYLATION.
CC DOMAIN 2 23 NEARLY ALL OF THE INHIBITORY ACTIVITY IS
CC PRESENT IN THIS REGION.
CC SEQUENCE 68 AA; 7707 MW; 8295BF7AF3E4A01A CRC64;
SQ

```

```

RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RL influenzae Rd."
RT Science 269:496-512(1995).
CC -1- FUNCTION: REQUIRED FOR PHAGE Q BETA RNA-DIRECTED SYNTHESIS OF
CC COMPLEMENTARY MINUS-STRAND RNA (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U32724; AAC2070.1; -
CC DR TIGR; HT0411; -
CC RNA-binding.
CC INIT_MER 0 0 BY SIMILARITY.
SQ SEQUENCE 90 AA; 10118 MW; 1C91D4ECF1F7F029 CRC64;

```

```

Query Match 25.0%; Score 6; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LRRERI 12
DB 14 LRRERI 19

RESULT 6
YBAC_SCHPO STANDARD; PRT; 93 AA.
AC 014358;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL 10.7 KDA PROTEIN C30D10.12C IN CHROMOSOME II.
GN SPC30D10.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID-4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Wood V., Rajandream M.A., Barrell B.G., Lauber J., Hilbert H.,
RL Duesterhoef A.;
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO YEAST YGR215W.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; Z97992; CAB10807.1; -
CC DR Hypothetical protein.
CC SEQUENCE 93 AA; 10698 MW; 6EC74F169DA747B CRC64;
SQ

```

```

Query Match 25.0%; Score 6; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 18 QLRLL 23
    111111
DB 55 QLRLL 60

RESULT 7
R17_ORISA
ID R17_ORISA STANDARD: PRT: 145 AA.
AC Q9ZST1;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE 30S RIBOSOMAL PROTEIN S17, CHLOROPLAST PRECURSOR (CS17).
GN RPS17.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
OC Oryza
OC NCBI_TaxID=4530;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-CV. ILPOOM; TISSUE-Leaf;
RA Lee J.-S., Eun M.-Y.;
RT "Molecular cloning and characterization of three nuclear-encoded
RT chloroplast precursor of ribosomal protein genes in rice.";
RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- SIMILARITY: BELONGS TO THE S17P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF095707; AAC64969.1; -.
DR HSSP: P23828; IRIP.
DR InterPro: IPR000266; -.
DR Pfam: PF00366; Ribosomal_S17; 1.
DR PRINTS: PR00973; RIBOSOMALS17.
DR KEGG: RIBOSOMAL PROTEIN; Chloroplast; Transit peptide.
FT TRANSIT 1
FT CHAIN ? 145 CHLOROPLAST PROTEIN S17.
SQ SEQUENCE 145 AA; 15975 MW; ABD46935DCE15E56 CRC64;

Query Match 25.0%; Score 6; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RRERIK 13
    111111
DB 68 RRERIK 73

RESULT 8
Y004_TREPA
ID Y004_TREPA STANDARD: PRT: 145 AA.
AC O83050;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHETICAL PROTEIN TP0004.
GN TP0004.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OC NCBI_TaxID=160;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;

```

```

RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artlach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE001185; AAC65005.1; -.
DR TIGR: TP0004;
RN Hypothetical protein.
SQ SEQUENCE 145 AA; 16338 MW; 6D6674C8BFFCFAC CRC64;

Query Match 25.0%; Score 6; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 EQLRTL 22
    111111
DB 130 EQLRTL 135

RESULT 9
LC4S_HUMAN
ID LC4S_HUMAN STANDARD: PRT: 150 AA.
AC Q16873;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LEUKOTRIENE C4 SYNTHASE (EC 2.5.1.37) (LEUKOTRIENE-C(4) SYNTHASE)
DE (LC4S SYNTHASE).
GN LC4S.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-22 AND 35-48.
RC TISSUE-Bone marrow;
RX MEDLINE=94329576; PubMed=8052639;
RA Lam B.K., Penrose J.F., Freeman G.J., Austen K.F.;
RT "Expression cloning of a cDNA for human leukotriene C4 synthase, an
RT integral membrane protein conjugating reduced glutathione to
RT leukotriene A4.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:7663-7667(1994).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=95024037; PubMed=7937884;
RA Welsh D.J., Creely D.P., Hauser S.D., Mathis K.J., Krivi G.G.,
RA Isakson P.C.;
RT "Molecular cloning and expression of human leukotriene-C4 synthase.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9745-9749(1994).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=96212205; PubMed=8626689;
RA Penrose J.F., Spector J., Baldasaro M., Xu K., Boyce J., Arm J.P.,
RA Austen K.F., Lam B.K.;
RT "Molecular cloning of the gene for human leukotriene C4 synthase.
RT Organization, nucleotide sequence, and chromosomal localization to
RT 5q35.";
RL J. Biol. Chem. 271:11356-11361(1996).

```



```

RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97054000; PubMed=8898379;
RA Bigby T.D., Hodulik C.R., Arden K.C., Fu L.;
RT "Molecular cloning of the human leukotriene C4 synthase gene and
  assignment to chromosome 5q35."
RL Mol. Med. 2:637-646(1996).
RN [5]
RP SEQUENCE OF 1-35.
RX MEDLINE=93189628; PubMed=8446623;
RA Nicholson D.W., Ali A., Vailancourt J.P., Calaycay J.R.,
  Mumford R.A., Zamoni R.J., Ford-Hutchinson A.W.;
RT "Purification to homogeneity and the N-terminal sequence of human
  leukotriene C4 synthase: a homodimeric glutathione S-transferase
  composed of 18-kDa subunits."
RL Proc. Natl. Acad. Sci. U.S.A. 90:2015-2019(1993).
CC -1- FUNCTION: CATALYZES THE CONJUGATION OF LEUKOTRIENE A4 WITH REDUCED
  GLUTATHIONE TO FORM LEUKOTRIENE C4.
CC -1- CATALYTIC ACTIVITY: LEUKOTRIENE A4 + GLUTATHIONE = LEUKOTRIENE C4
  + H(2)O.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: LTC4S ACTIVITY IS PRESENT IN EOSINOPHILS,
  BASOPHILS, MAST CELLS, CERTAIN PHAGOCYtic MONONUCLEAR CELLS,
  ENDOTHELIAL CELLS, VASCULAR SMOOTH MUSCLE CELLS, AND PLATELETS.
CC -1- SIMILARITY: BELONGS TO THE MAPBG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U09353; AAA20467.1; -
DR EMBL: U11552; AAA50555.1; -
DR EMBL: U50136; AAC50476.1; -
DR EMBL: U62025; AAB06723.1; -
DR MIM: 246530; -
DR InterPro: IPR001129; -
DR Pfam: PF01124; MAPEG; 1.
DR PROSITE: PS01297; FLAP_GST2_LTC4S; 1.
KM Transferrase; Transmembrane; Leukotriene biosynthesis.
FT TRANSMEM 7 27
FT TRANSMEM 68 88
FT TRANSMEM 115 135
FT CONFLICT 21 21 Y -> G (IN REF. 5).
SQ SEQUENCE 150 AA; 16566 MW; 04E269B475063037 CRC64;

Query Match 25.0%; Score 6; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LRTLLP 24
DB 143 LRTLLP 148

RESULT 10
GREA_BUCAI
ID GREA_BUCAI STANDARD; PRT; 159 AA.
AC P57464;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TRANSCRIPTION ELONGATION FACTOR GREA (TRANSCRIPT CLEAVAGE FACTOR
  GREA).
GN GREA OR BU384.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
  symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.

```

```

OX NCBI_TaxId=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
  Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -1- FUNCTION: NECESSARY FOR EFFICIENT RNA POLYMERASE TRANSCRIPTION
  OF ELONGATION PAST TEMPLATE-ENCODED ARRESTING SITES. THE ARRESTING
  SITES IN DNA HAVE THE PROPERTY OF TRAPPING A CERTAIN FRACtion
  OF ELONGATING RNA POLYMERASES THAT PASS THROUGH, RESULTING IN
  LOCKED TERNARY COMPLEXES. CLEAVAGE OF THE NASCENT TRANSCRIPT BY
  CLEAVAGE FACTORS SUCH AS GREA OR GREB ALLOWS THE RESUMPTION OF
  ELONGATION FROM THE NEW 3'TERMINUS. GREA RELEASES SEQUENCES OF
  2 TO 3 NUCLEOTIDES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GREA/GREB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP001119; BAB13087.1; -
DR PROSITE: PS00829; GREAB.1; 1.
DR PROSITE: PS00830; GREAB.2; 1.
KM Transcription regulation; DNA-binding; Coiled coil.
FT DOMAIN 10 27
SQ SEQUENCE 159 AA; 17983 MW; A400D7A95EFD1D3 CRC64;

Query Match 25.0%; Score 6; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 EKLRRE 10
DB 13 EKLRRE 18

RESULT 11
YMT0_YEAST
ID YMT0_YEAST STANDARD; PRT; 160 AA.
AC C004210;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 19.2 KDA PROTEIN IN SOB1-ARGL1 INTERGENIC REGION.
GN YMR040W OR YM9532.05.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
  Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Odell C., Bowman S., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: TO YEAST YKL065C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: Z48502; CA88406.1; -  
 DR SGD; S0004643; YMR040W.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 3 23 POTENTIAL.  
 FT TRANSMEM 46 66 POTENTIAL.  
 FT TRANSMEM 104 124 POTENTIAL.  
 SQ SEQUENCE 160 AA; 19178 MW; D003EC7412559871 CRC64;

Query Match 25.0%; Score 6; DB 1; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKLR 9  
 DB 142 KEKLR 147

RESULT 12  
 ID PUR6\_PYRHO STANDARD; PRT; 177 AA.  
 AC 058058;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PHOSPHORIBOSYLAMINOIMIDAZOLE CARBOXYLASE CATALYTIC SUBUNIT  
 DE (EC 4.1.1.21) (AIR CARBOXYLASE) (AIRC).  
 GN PURE OR PH0320 OR PHCD015.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 NC NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE=98344137; PubMed=9679194;  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohiku Y.,  
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 CC -1- FUNCTION: THIS SUBUNIT CAN ALONE TRANSFORM AIR TO CAIR, BUT IN  
 CC ASSOCIATION WITH PURK, WHICH POSSESSES AN ATPASE ACTIVITY, AN  
 CC ENZYME COMPLEX IS PRODUCED WHICH IS CAPABLE OF CONVERTING AIR TO  
 CC CAIR EFFICIENTLY UNDER PHYSIOLOGICAL CONDITION (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 1-(5-PHOSPHORIBOSYL)-5-AMINO-4-IMIDAZOLE-  
 CC CARBOXYLATE - 1-(5-PHOSPHORIBOSYL)-5-AMINOIMIDAZOLE + CO(2).  
 CC -1- PATHWAY: SIXTH STEP IN DE NOVO PURINE BIOSYNTHESIS.  
 CC -1- SIMILARITY: TO OTHER BACTERIAL PURE, ALSO TO AIR CARBOXYLASE FROM  
 CC FUNGI.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AP000002; BAA29394.1; -  
 DR InterPro: IPR000031; -  
 DR Pfam: PF00731; AIRC; 1.  
 KW Purine biosynthesis; Lyase; Decarboxylase;  
 SQ SEQUENCE 177 AA; 19439 MW; F86848BA655703F CRC64;

Query Match 25.0%; Score 6; DB 1; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKLR 9  
 DB 145 KEKLR 150

RESULT 13  
 ID NOHB\_ECOLI STANDARD; PRT; 181 AA.  
 AC P31062; P75721;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE PROPHAGE QSR DNA PACKAGING PROTEIN NOI HOMOLOG.  
 GN NOHB.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 NC NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [2]

RP SEQUENCE OF 1-147 FROM N.A.  
 RC STRAIN=K12 / W3110;  
 RX MEDLINE=92334985; PubMed=1630906;  
 RA Kotani H., Kawamura A., Takahashi A., Nakatsuji M., Hiraoka N.,  
 RA Nakajima K., Takamami M.;  
 RT "Site-specific dissection of E. coli chromosome by lambda terminase.";  
 RL Nucleic Acids Res. 20:3357-3360(1992).  
 CC -1- SIMILARITY: TO PHAGE LAMBDA DNA PACKAGING PROTEIN NUI.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AE000161; AAC73661.1; -  
 DR EMBL: D00928; -; NOT\_ANNOTATED\_CDS.  
 DR Ecogene: EG1635; nohb.  
 SQ SEQUENCE 181 AA; 20428 MW; E38D4FEFF174751C CRC64;

Query Match 25.0%; Score 6; DB 1; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 KEKLR 10  
 DB 60 KEKLR 65

RESULT 14  
 ID TERS\_LAMB STANDARD; PRT; 181 AA.  
 AC P03707;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE TERMINASE SMALL SUBUNIT (DNA PACKAGING PROTEIN NUI).  
 GN NUI.  
 OS Bacteriophage lambda.  
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;  
 OC Lambda phage group.

```

OX NCBI_TaxID=10710;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=83189071; PubMed=6221115;
RA Sanger F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B.;
RT "Nucleotide sequence of bacteriophage lambda DNA.";
RL J. Mol. Biol. 162:729-773(1982).
RN [2]
RN ATP-BINDING DOMAIN.
RX MEDLINE=88172462; PubMed=2965248;
RA Becker A., Gold M.;
RT "Prediction of an ATP reactive center in the small subunit, gpnu1, of
the phage lambda terminase enzyme.";
RL J. Mol. Biol. 199:219-222(1988).
CC -1- FUNCTION: INVOLVED IN THE INITIATION OF LAMBDA DNA PACKAGING INTO
THE PROHEAD. PROCESSES REPLICATING CONCATEMERIC DNA INTO PIECES
OF UNIT LENGTH WITH COHESIVE ENDS.
CC -1- SUBUNIT: HETEROOLIGOMER OF NU1 AND GPA.
CC -1- SIMILARITY: TO BACTERIOPHAGE 21 TERMINASE SMALL SUBUNIT (GP1).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: J02459; AAA96533.1; -.
DR PIR: A04329; JVBPNL.
KW DNA packaging; ATP-binding.
FT NP_BIND 31 36
SQ SEQUENCE 181 AA; 20441 MW; A25A5EAEF7BF18A4 CRC64;

```

```

Query Match          25.0%; Score 6; DB 1; Length 181;
Best Local Similarity 100.0%; Pred. NO. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 5 EKLRE 10
    |||||
DB 60 EKLRE 65

```

```

RESULT 15
NOHA_ECOLI STANDARD; PRT; 189 AA.
ID P31061; P77152;
AC 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROPHAGE QIN DNA PACKAGING PROTEIN NU1 HOMOLOG.
GN NOHA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,

```

```

RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sempel G., Seki Y., Sivasubaram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horituchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28,0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RN SEQUENCE OF 1-147 FROM N.A.
RP STRAIN=K12 / W3110;
RX MEDLINE=92334985; PubMed=1630906;
RA Kotani H., Kawamura A., Takahashi A., Nakatsuji M., Hiraoka N.,
RA Nakajima K., Takamami M.;
RT "Site-specific dissection of E. coli chromosome by lambda terminase.";
RL Nucleic Acids Res. 20:3357-3360(1992)
CC -1- SIMILARITY: TO PHAGE LAMBDA DNA PACKAGING PROTEIN NU1.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AE000252; AAC74621.1; -.
DR EMBL: D90798; BA15252.1; -.
DR EMBL: D00927; -. NOT_ANNOTATED_CDS.
DR EcoGene: EG11634; noha.
SQ SEQUENCE 189 AA; 21404 MW; FB5B1F844AB2C7C1 CRC64;

```

```

Query Match          25.0%; Score 6; DB 1; Length 189;
Best Local Similarity 100.0%; Pred. NO. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 5 EKLRE 10
    |||||
DB 60 EKLRE 65

```

Search completed: September 15, 2001, 12:50:40  
Job time: 285 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:50:05 : Search time 69.13 Seconds  
(without alignments)  
45.933 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_163

Perfect score: 24

Sequence: 1 HSKSEKLRERIKYCCQLRTLPLP 24

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_16:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.trodet:\*
- 12: sp.unclassified:\*
- 13: sp.vertebrate:\*
- 14: sp.virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	75.0	425	4	Q9NX45
2	18	33.3	1220	10	Q24015
3	8	33.3	1240	10	Q24016
4	8	33.3	1266	10	Q24017
5	7	29.2	1250	10	Q9XET3
6	7	29.2	331	14	Q9MOV5
7	7	29.2	443	10	Q11316
8	7	29.2	443	10	Q9LYR0
9	7	29.2	502	2	Q9ZB31
10	7	29.2	746	10	Q9FNF6
11	7	29.2	748	14	Q98200
12	7	29.2	1080	11	Q9WVP5
13	7	29.2	1080	11	Q9E0U1
14	7	29.2	1260	5	Q9GRG5
15	7	29.2	1341	3	Q08281
16	7	29.2	1342	3	Q08281
17	7	29.2	1833	5	Q9VM67
18	6	25.0	39	4	Q9UBN9
19	6	25.0	64	2	Q9EY25

20	6	25.0	81	2	Q52484	Q52484 burkholderi
21	6	25.0	87	2	Q9KV11	Q9KV11 vibrio chol
22	6	25.0	88	14	Q9EAW8	Q9EAW8 human immun
23	6	25.0	90	1	Q26558	Q26558 methanobact
24	6	25.0	94	1	Q9HH63	Q9HH63 methanobact
25	6	25.0	94	9	Q80193	Q80193 methanobact
26	6	25.0	109	2	Q9KPE0	Q9KPE0 vibrio chol
27	6	25.0	131	10	Q9XES9	Q9XES9 glycine max
28	6	25.0	132	8	Q9TJN6	Q9TJN6 zea mays (m
29	6	25.0	135	5	Q17714	Q17714 caenorhabdi
30	6	25.0	143	2	Q9L9G4	Q9L9G4 streptomyce
31	6	25.0	160	5	Q93970	Q93970 caenorhabdi
32	6	25.0	171	10	Q9FR14	Q9FR14 brassica na
33	6	25.0	184	3	Q13954	Q13954 schizosacch
34	6	25.0	185	2	Q9F2S8	Q9F2S8 streptomyce
35	6	25.0	201	2	Q9RR40	Q9RR40 citrobacter
36	6	25.0	211	2	Q9RW74	Q9RW74 deinococcus
37	6	25.0	215	2	Q915R6	Q915R6 pseudomonas
38	6	25.0	218	2	Q9RUE2	Q9RUE2 deinococcus
39	6	25.0	218	5	Q9V4H1	Q9V4H1 drosophila
40	6	25.0	219	1	Q58180	Q58180 pyrococcus
41	6	25.0	222	2	Q9WY98	Q9WY98 thermotoga
42	6	25.0	222	10	Q38918	Q38918 arabidopsis
43	6	25.0	232	2	Q9RRF3	Q9RRF3 deinococcus
44	6	25.0	239	2	Q9K744	Q9K744 bacillus ha
45	6	25.0	242	1	Q9UXY9	Q9UXY9 pyrococcus

## ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	425 AA.
Q9NX45				
ID	Q9NX45			
AC	Q9NX45			
DT	01-OCT-2000 (TREMBlrel. 15, Created)			
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)			
DT	01-MAR-2001 (TREMBlrel. 16, Last annotation update)			
DE	CDNA FLJ20049 FIS, CLONE KAT05575.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,			
RA	Suzuki Y., Ohtsuka M., Nishitani T., Shibahara T., Tanaka T.,			
RA	Nakamura Y., Isono T., Sugano S.,			
RT	"NPO human cDNA sequencing project."			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AK000456; BAA91175.1; -			
DR	InterPro: IPR001092; -			
DR	SMART; SM00353; HLH; 1.			
DR	SEQUENCE 425 AA; 46922 MW; 777E57231F9D6BDD CRC64;			
QY	7 LRRERIKYCCQLRTLPLP 24			
DB	212 LRRERIKYCCQLRTLPLP 229			
RESULT	2			
ID	Q24015			
AC	Q24015			
DT	01-JAN-1998 (TREMBlrel. 05, Created)			
DT	01-JAN-1998 (TREMBlrel. 05, Last sequence update)			
DT	01-MAR-2001 (TREMBlrel. 16, Last annotation update)			
DE	RESISTANCE COMPLEX PROTEIN I2C-1.			

GN 12C-1.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
 OC Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97290204; PubMed=9144960;  
 RA Ori N., Eshed Y., Paran I., Presting G., Aviv D., Tanksley S.,  
 RA Zamir D., Fluhr R.;  
 RT "The 12C family from the wilt disease resistance locus I2 belongs to  
 RT the nucleotide binding, leucine-rich repeat superfamily of plant  
 RT resistance genes.";  
 RL Plant Cell 9:521-532(1997).  
 DR EMBL: AF004878; AAB63274.1; -;  
 DR Mendel: 24472; Lyces:3172;24472.  
 DR InterPro: IPR000767; -;  
 DR InterPro: IPR001611; -;  
 DR InterPro: IPR002182; -;  
 DR InterPro: IPR003592; -;  
 DR Pfam: PF00560; LRR; 4.  
 DR Pfam: PF00931; NB-ARC; 1.  
 DR PRINTS: PR00364; DISEASERIST.  
 DR SMART: SM00370; LRR; 1.  
 SQ SEQUENCE 1220 AA; 139386 MW; B5997280C8C5EB20 CRC64;

Query Match 33.3%; Score 8; DB 10; Length 1220;  
 Best Local Similarity 100.0%; Pred. No. 7.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 EOLRTLLP 24  
 |||||  
 DB 554 EOLRTLLP 561

RESULT 3  
 024016 PRELIMINARY; PRT; 1240 AA.  
 AC 024016;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE RESISTANCE COMPLEX PROTEIN 12C-2.  
 GN 12C-2.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
 OC Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97290204; PubMed=9144960;  
 RA Ori N., Eshed Y., Paran I., Presting G., Aviv D., Tanksley S.,  
 RA Zamir D., Fluhr R.;  
 RT "The 12C family from the wilt disease resistance locus I2 belongs to  
 RT the nucleotide binding, leucine-rich repeat superfamily of plant  
 RT resistance genes.";  
 RL Plant Cell 9:521-532(1997).  
 DR EMBL: AF004879; AAB63275.1; -;  
 DR Mendel: 24473; Lyces:3172;24473.  
 DR InterPro: IPR001611; -;  
 DR InterPro: IPR002182; -;  
 DR InterPro: IPR003592; -;  
 DR Pfam: PF00560; LRR; 4.  
 DR Pfam: PF00931; NB-ARC; 1.  
 DR SMART: SM00370; LRR; 1.  
 SQ SEQUENCE 1240 AA; 141523 MW; D74EB158C406C102 CRC64;

Query Match 33.3%; Score 8; DB 10; Length 1240;  
 Best Local Similarity 100.0%; Pred. No. 7.4;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 17 EOLRTLLP 24  
 |||||  
 DB 545 EOLRTLLP 552.

RESULT 4  
 09XET3 PRELIMINARY; PRT; 1266 AA.  
 AC 09XET3;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE DISEASE RESISTANCE PROTEIN I2.  
 GN I2.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
 OC Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98299845; PubMed=9634592;  
 RA Simons G., Groenendijk J., Wijnbrandi J., Reijans M., Groenen J.,  
 RA Diergaarde P., Van der Lee T., Bleeker M., Onstienk J., De Both M.,  
 RA Haring M., Mes J., Cornelissen B., Zabeau M., Vos P.;  
 RT "Dissection of the fusarium I2 gene cluster in tomato reveals six  
 RT homologs and one active gene copy.";  
 RL Plant Cell 10:1055-1068(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Simons G., Groenendijk J., Wijnbrandi J., Reijans M., Groenen J.,  
 RA Diergaarde P., Van der Lee T., Bleeker M., Onstienk J., De Both M.,  
 RA Haring M., Mes J., Cornelissen B., Zabeau M., Vos P.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF118127; AAD27815.1; -;  
 DR InterPro: IPR000767; -;  
 DR InterPro: IPR001593; -;  
 DR InterPro: IPR001611; -;  
 DR InterPro: IPR002182; -;  
 DR Pfam: PF00560; LRR; 3.  
 DR Pfam: PF00931; NB-ARC; 1.  
 DR PRINTS: PR00364; DISEASERIST.  
 DR Prodom: PD003035; -; 1.  
 SQ SEQUENCE 1266 AA; 144826 MW; 8392FFBE7FFD7F5D CRC64;

Query Match 33.3%; Score 8; DB 10; Length 1266;  
 Best Local Similarity 100.0%; Pred. No. 7.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 EOLRTLLP 24  
 |||||  
 DB 547 EOLRTLLP 554

RESULT 5  
 09M0V5 PRELIMINARY; PRT; 250 AA.  
 AC 09M0V5;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE HYPOHETICAL 27.6 KDA PROTEIN.  
 GN AT4603400.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosidids II;  
 OC Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,  
 RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,  
 RA Shekher M., Schuit K., See L.H., Swaby I., Habermann K., Dedhia N.N.,  
 RA Mewes H.W., Lemcke K., Mayer K.F.X.,  
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RP [12]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL161503; CAB81082.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 250 AA; 27627 MW; 7AAA04PBF0A6F707 CRC64;

Query Match 29.2%; Score 7; DB 10; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSKERLR 8  
 DB 244 SSKERLR 250

RESULT 6  
 ID 011316 PRELIMINARY; PRT; 331 AA.  
 AC 011316:  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
 DE CLONE B1-13 HOMOLOG OF VACCINIA E2L (B1-13) (FRAGMENT).  
 GN B1-13.  
 OS Molluscum contagiosum virus subtype 1 (MCV1).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Molluscipoxvirus.  
 OX NCBI\_TaxID=10280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,  
 RA Lopez-Esteban J.L., Esteban M., Martin-Gallardo A.,  
 RL Virus Genes 0:0-0(0).  
 DR EMBL: U86899; AAB57946.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 331 331  
 SQ SEQUENCE 331 AA; 37009 MW; E776B39950B997E CRC64;

Query Match 29.2%; Score 7; DB 14; Length 331;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 YCCEQLR 20  
 DB 127 YCCEQLR 133

RESULT 7  
 ID 09LYR0 PRELIMINARY; PRT; 443 AA.  
 AC 09LYR0:  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE HYPOTHETICAL 50.8 KDA PROTEIN.  
 GN T22N19\_110.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,

RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;  
 RN Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RP [12]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL163572; CAB87153.1; -.  
 DR InterPro: IPR000048; -.  
 DR Pfam: PF00642; IQ; 2.  
 DR SMART: SM00015; IQ; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 443 AA; 50820 MW; EA3C8B5E2041D83 CRC64;

Query Match 29.2%; Score 7; DB 10; Length 443;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LRRERIK 13  
 DB 227 LRRERIK 233

RESULT 8  
 ID 09ZB31 PRELIMINARY; PRT; 502 AA.  
 AC 09ZB31:  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE ERYTHRITOL PHOSPHATE DEHYDROGENASE.  
 GN ERIB.  
 OS Brucella abortus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=235;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=2308;  
 RA Sangari F.J., Aguerro J., Garcia-Ibobo J.M.;  
 RT "Characterization of a gene cluster involved in erythritol catabolism  
 in Brucella abortus.";  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U57100; AAD11520.1; -.  
 DR InterPro: IPR000447; -.  
 DR Pfam: PF01224; FAD\_Gly3P\_dh; 1.  
 DR PRINTS: PR01001; FADG3PDH.  
 DR PROSITE: PS00977; FAD\_G3PDH\_1; 1.  
 SQ SEQUENCE 502 AA; 56231 MW; A88459636FAD7A11 CRC64;

Query Match 29.2%; Score 7; DB 2; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEXLRRE 10  
 DB 307 KEXLRRE 313

RESULT 9  
 ID 09FN6 PRELIMINARY; PRT; 746 AA.  
 AC 09FN6:  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE DIPEPTIDYL PEPTIDASE IV-LIKE PROTEIN.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE-98069011; PubMed-9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT DNA Res. 4:291-300(1997).";
RL EMBL: AB06701; BAB10391.1; -
DR EMBL: AB06701; BAB10391.1; -
SQ SEQUENCE 746 AA; 83851 MW; AEC8ED567BEEB0A5 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EKLRRER 11
DB 80 EKLRRER 86

RESULT 10
Q98200 PRELIMINARY; PRT; 748 AA.
AC Q98200;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE MC032L.
GN MC032L.
OS Molluscum contagiosum virus subtype 1 (MCV1).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
OX NCBI_TaxID=10280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96325459; PubMed-8670425;
RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
RA Moss B.;
RT "Genome sequence of a human tumorigenic poxvirus: prediction of
RT specific host response-evasion genes.";
RL Science 273:813-816(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
RA Moss B.;
RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: U60315; AAC55160.1; -
SQ SEQUENCE 748 AA; 84411 MW; 61A2741F62CA5BA CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 YCCEQLR 20
DB 519 YCCEQLR 525

RESULT 11
Q45358 PRELIMINARY; PRT; 791 AA.
AC Q45358; Q45404;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE F26D2.2 PROTEIN.
GN F26D2.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
```

```

OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurtry A.;
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-94150718; PubMed-7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurtry A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierley-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Wellstock L., Wilkinson-Sproat J., Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA McMurtry A.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: Z93377; CAB07581.1; -
DR EMBL: Z81513; CAB07581.1; JOINED.
DR EMBL: Z81513; CAB04186.1; -
DR EMBL: Z93377; CAB04186.1; JOINED.
SQ SEQUENCE 791 AA; 91505 MW; 9B53803A9B6363F0 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EKLRRER 11
DB 153 EKLRRER 159

RESULT 12
Q9WVP5 PRELIMINARY; PRT; 1080 AA.
AC Q9WVP5;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PUTATIVE DEUBIQUITINATING ENZYME UBPY.
GN UBPY.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Martegani E., Gnesutta N., Mauri I., Bini V.;
RL Submitted (Apr-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF057146; AAD38869.1; -
DR MEROPS: C19.011; -
DR InterPro: IPR001394; -
DR InterPro: IPR001448; -
DR InterPro: IPR001763; -
DR Pfam: PF00442; UCH-1; 1.
DR Pfam: PF00443; UCH-2; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR PROSITE: PS00304; SASP_1; UNKNOWN_1.
DR PROSITE: PS00972; UCH_2_1; 1.
DR PROSITE: PS00973; UCH_2_2; 1.
DR PROSITE: PS00973; UCH_2_3; 1.
DR PROSITE: PS00973; UCH_2_3; 1.
DR SMART: SM00450; RHOD; 1.
```



SQL SEQUENCE 1080 AA; 122578 MW; 2BF00AFF68AFC8B5 CRC64;

Query Match 29.2%; Score 7; DB 11; Length 1080;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKLRRE 10  
|||||||  
DB 485 KEKLRRE 491

RESULT 13

O9EQU1 PRELIMINARY; PRT; 1080 AA.  
AC O9EQU1.  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE DEUBICUTINATING ENZYME UBPY.  
GN MUBPY.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=20545486; PubMed=10982817;  
RA Kato M., Miyazawa K., Kitamura N.;  
RA "A Deubiquitinating Enzyme UBPY Interacts with the Src Homology 3  
RT Domain of Hrs-binding Protein via a Novel Binding Motif  
RT PX(V/I)(D/N)RXKP."  
RL J. Biol. Chem. 275:37481-37487(2000).  
DR EMBL, AB045709; BAB18534.1; -;  
SQ SEQUENCE 1080 AA; 122460 MW; C1B870FF98C3224 CRC64;

Query Match 29.2%; Score 7; DB 11; Length 1080;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKLRRE 10  
|||||||  
DB 485 KEKLRRE 491

RESULT 14  
O9GRG5 PRELIMINARY; PRT; 1260 AA.

AC O9GRG5;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE PUTATIVE STRUCTURAL MAINTENANCE OF CHROMOSOME 3 PROTEIN.  
GN PUTATIVE SMC3.  
OS Trypanosoma brucei.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxID=5691;  
RN [1]  
RN SEQUENCE FROM N.A.  
RA Ersfeld K., Walgraffe D., Gull K.;  
RA "Identification of an SMC3-homologue in Trypanosoma brucei."  
RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL, AJ279087; CAC12695.1; -;  
SQ SEQUENCE 1260 AA; 143958 MW; 2074957B2146DE06 CRC64;

DB 887 KEKLRRE 893

RESULT 15

O08281 PRELIMINARY; PRT; 1341 AA.  
AC O08281;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE CHROMOSOME XV READING FRAME ORF YOL138C.  
GN YOL138C.  
OS Saccharomyces cerevisiae (Baker's Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RN SEQUENCE FROM N.A.  
RA Arino J., Casamayor A., Gamo F.J., Gancedo C., Lafuente M.J.,  
RA Aldea M., Casas C., Herrero E.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RN SEQUENCE FROM N.A.  
RA MIPS;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS A RING (C3HC4-CLASS) ZINC FINGER.  
DR EMBL, Z74880; CAA9159.1; -;  
DR SGD, S0005498; YOL138C.  
DR InterPro: IPR001680; -;  
DR InterPro: IPR001841; -;  
DR InterPro: IPR002052; -;  
DR Pfam: PR00097; zf-C3HC4; 1.  
DR Pfam: PR00400; WD40; 2.  
DR PROSITE: PS00092; N6\_MTASE; UNKNOWN\_1.  
DR PROSITE: PS00678; WD\_REPEATS; UNKNOWN\_1.  
DR SMART: SM00184; RING; 1.  
KW Repeat; WD repeat; Zinc-finger.  
SQ SEQUENCE 1341 AA; 149268 MW; EB0EAD034818EA0 CRC64;

Query Match 29.2%; Score 7; DB 3; Length 1341;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSKEKLR 8  
|||||||  
DB 1257 SSKEKLR 1263

Search completed: September 15, 2001, 12:50:11  
Job time: 271 sec

Query Match 29.2%; Score 7; DB 5; Length 1260;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KEKLRRE 10  
|||||||

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:45:40 ; Search time 69.13 Seconds  
(without alignments)  
32.536 Million cell updates/sec

Title: us-09-389-000-2\_copy\_134\_150  
Perfect score: 17  
Sequence: 1 KRISLHSSKEKLRRER 17

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

SPTRMBL\_16:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.protoent:\*  
12: sp.unclassified:\*  
13: sp.vertebrate:\*  
14: sp.virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	64.7	425	4 Q9NX45	Q9NX45 homo sapien
2	7	41.2	250	4 Q9MOV5	Q9MOV5 arabidopsis
3	7	41.2	416	4 Q9H533	Q9H533 homo sapien
4	7	41.2	502	2 Q9ZB31	Q9ZB31 brucea ab
5	7	41.2	549	8 Q9T6Y9	Q9T6Y9 spathiphyll
6	7	41.2	559	8 Q9T6B9	Q9T6B9 ascarina sp
7	7	41.2	561	8 Q9T6T8	Q9T6T8 doryphora s
8	7	41.2	562	8 Q9T6T6	Q9T6T6 pennus boid
9	7	41.2	562	8 Q9T6T0	Q9T6T0 akedia quin
10	7	41.2	565	8 Q9T6U6	Q9T6U6 trochodendr
11	7	41.2	567	8 Q9T701	Q9T701 thotica tom
12	7	41.2	577	8 Q9T6W3	Q9T6W3 lirioidendro
13	7	41.2	579	8 Q9T6V8	Q9T6V8 idiospermum
14	7	41.2	580	8 Q9T6V1	Q9T6V1 eupatelia po
15	7	41.2	581	8 Q9T6X3	Q9T6X3 podophyllum
16	7	41.2	584	8 Q9T6X3	Q9T6X3 polyalthia
17	7	41.2	586	8 Q9T6X8	Q9T6X8 cinnaomend
18	7	41.2	586	8 Q9T6X0	Q9T6X0 knema later
19	7	41.2	586	8 Q9T6W1	Q9T6W1 cocculus tr

## ALIGNMENTS

Result ID	Query Match	Score	Length	DB ID	Description
1	64.7%	Score 11; DB 4; Length 425;			
Q9NX45	Best Local Similarity 100.0%; Pred. No. 0.004;				
Q9NX45	Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Q9NX45	1 KRISLHSSKE 11				
Q9NX45	200 KRISLHSSKE 210				
Q9MOV5	1 KRISLHSSKE 11				
Q9MOV5	200 KRISLHSSKE 210				
Q9H533	1 KRISLHSSKE 11				
Q9H533	200 KRISLHSSKE 210				
Q9ZB31	1 KRISLHSSKE 11				
Q9ZB31	200 KRISLHSSKE 210				
Q9T6Y9	1 KRISLHSSKE 11				
Q9T6Y9	200 KRISLHSSKE 210				
Q9T6B9	1 KRISLHSSKE 11				
Q9T6B9	200 KRISLHSSKE 210				
Q9T6T8	1 KRISLHSSKE 11				
Q9T6T8	200 KRISLHSSKE 210				
Q9T6T6	1 KRISLHSSKE 11				
Q9T6T6	200 KRISLHSSKE 210				
Q9T6U6	1 KRISLHSSKE 11				
Q9T6U6	200 KRISLHSSKE 210				
Q9T701	1 KRISLHSSKE 11				
Q9T701	200 KRISLHSSKE 210				
Q9T6W3	1 KRISLHSSKE 11				
Q9T6W3	200 KRISLHSSKE 210				
Q9T6V8	1 KRISLHSSKE 11				
Q9T6V8	200 KRISLHSSKE 210				
Q9T6V1	1 KRISLHSSKE 11				
Q9T6V1	200 KRISLHSSKE 210				
Q9T6X3	1 KRISLHSSKE 11				
Q9T6X3	200 KRISLHSSKE 210				
Q9T6X8	1 KRISLHSSKE 11				
Q9T6X8	200 KRISLHSSKE 210				
Q9T6X0	1 KRISLHSSKE 11				
Q9T6X0	200 KRISLHSSKE 210				
Q9T6W1	1 KRISLHSSKE 11				
Q9T6W1	200 KRISLHSSKE 210				

GN ATAG05400.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Spiegel L.A., Huang E.N., Nascimento L.V., de la Bastide M., Vil D.M.,  
 RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,  
 RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,  
 RA Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AL161503; CAB81082.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 250 AA; 27627 MW; 7AANA04BF0A6F707 CRC64;

Query Match 41.2%; Score 7; DB 10; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SSKERLR 14  
 DB 244 SSKERLR 250

RESULT 3  
 OYH553 PRELIMINARY; PRT; 416 AA.  
 AC OYH553;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE BA13B9.1 (NOVEL PROTEIN SIMILAR TO A GLYCOSYLTRANSFERASE).  
 GN BA13B9.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Babage A.;  
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AL137067; CAC07999.1; -.  
 KW Transferase.  
 SQ SEQUENCE 416 AA; 47091 MW; 778DB1FD069E7F29 CRC64;

Query Match 41.2%; Score 7; DB 4; Length 416;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KISLHLS 8  
 DB 308 KISLHLS 314

RESULT 4  
 OYB31 PRELIMINARY; PRT; 502 AA.  
 AC OYB31;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE ERYTHRITOL PHOSPHATE DEHYDROGENASE.  
 GN ERIB.  
 OS Brucella abortus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=235;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-2308;  
 RA Sangari F.J., Aguero J., Garcia-Lobo J.M.;  
 RT "Characterization of a gene cluster involved in erythritol catabolism  
 RT in Brucella abortus.";  
 RL Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: U57100; A011520.1; -.  
 DR InterPro: IPR000447; -.  
 DR Pfam: PF01224; FAD\_Gly3P\_dh; 1.  
 DR PRINTS: PR01001; FADG3PDH.  
 DR PROSITE: PS00977; FAD\_G3PDH\_1; 1.  
 SQ SEQUENCE 502 AA; 56231 MW; A88459636F4D7A11 CRC64;

Query Match 41.2%; Score 7; DB 2; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 KEKLRRE 16  
 DB 307 KEKLRRE 313

RESULT 5  
 OY6Y9 PRELIMINARY; PRT; 549 AA.  
 AC OY6Y9;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE MATURASE (FRAGMENT).  
 GN MATR.  
 OS Spathiphyllum clevealandii.  
 OC Mitochondrion.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Araceae; Spathiphyllum.  
 OX NCBI\_TaxID=78382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20052162; PubMed-10586879;  
 RA Qiu Y.-L., Lee J., Bernasconi-Quadrone F., Solits D.E., Solits P.S.,  
 RA Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;  
 RT "The Earliest Angiosperms: Evidence from Mitochondrial, Plastid and  
 RT Nuclear Genomes."  
 RL Nature 402:404-407(1999).  
 CC -I- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE  
 CC TRANSCRIPTASE).  
 DR EMBL: AF197746; AAF14738.1; -.  
 DR InterPro: IPR000442; -.  
 DR InterPro: IPR000477; -.  
 DR Pfam: PF00078; rvt; 2.  
 DR Pfam: PF01348; Intron\_maturase2; 1.  
 KW Mitochondrion; RNA-directed DNA polymerase.  
 FT NON\_TER 1  
 FT NON\_TER 549  
 SQ SEQUENCE 549 AA; 61623 MW; 1C3B8E573A699DEC CRC64;

Query Match 41.2%; Score 7; DB 8; Length 549;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9  
 DB 427 ISLHSS 433

RESULT 6  
 OY6B9 PRELIMINARY; PRT; 559 AA.  
 ID OY6B9

```

AC 09t6b9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE MATURASE (FRAGMENT).
GN MATR.
OS Ascarina sp. Qiu-M149.
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Chloranthaceae; Ascarina.
OX NCBI_TaxID=108576;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20052162; PubMed=10586879;
RA Qiu Y.-L., Lee J., Bernasconi-Quadrone F., Soltis D.E., Soltis P.S.,
RA Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;
RT "The Earliest Angiosperms: Evidence from Mitochondrial, Plastid and
RT Nuclear Genomes.";
RL Nature 402:404-407(1999).
CC -I- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC GRADIENT ACROSS THE MEMBRANE. THE ALPHA CHAIN IS A REGULATOR
CC SUBUNIT (BY SIMILARITY).
CC -I- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C (BY SIMILARITY).
CC -I- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
CC TRANSCRIPTASE).
CC -I- SIMILARITY: TO ATP SYNTHASE ALPHA CHAIN, C TERMINUS.
DR EMBL; AF197755; AAF14747.1; -
DR InterPro; IPR000442; -
DR InterPro; IPR000477; -
DR InterPro; IPR000790; -
DR Pfam; PF00078; rvt; 2.
DR Pfam; PF01348; Intron_maturas2; 1.
DR ProDom; PD001099; -; 1.
KW ATP synthetase; ATP-binding; CF(1); Hydrogen ion transport; Hydrolase;
KW Mitochondrion; RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 559 AA; 62857 MW; E343B0A3E976255F CRC64;

Query Match 41.2%; Score 7; DB 8; Length 559;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9
DB 441 ISLHSS 447

RESULT 7
O9t6T8 PRELIMINARY; PRT; 561 AA.
AC O9t6T8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE MATURASE (FRAGMENT).
GN MATR.
OS Doryphora sassafras.
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Laurales; Monimiacae; Doryphora.
OX NCBI_TaxID=74884;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20052162; PubMed=10586879;
RA Qiu Y.-L., Lee J., Bernasconi-Quadrone F., Soltis D.E., Soltis P.S.,
RA Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;
RT "The Earliest Angiosperms: Evidence from Mitochondrial, Plastid and
RT Nuclear Genomes.";

```

```

RL Nature 402:404-407(1999).
CC -I- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
CC TRANSCRIPTASE).
DR EMBL; AF197801; AAF14793.1; -
DR InterPro; IPR000442; -
DR InterPro; IPR000477; -
DR Pfam; PF00078; rvt; 1.
DR Pfam; PF01348; Intron_maturas2; 1.
KW Mitochondrion; RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 561 AA; 62921 MW; 51A8F9F8F2C323D7 CRC64;

Query Match 41.2%; Score 7; DB 8; Length 561;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9
DB 443 ISLHSS 449

RESULT 8
O9t6T6 PRELIMINARY; PRT; 562 AA.
AC O9t6T6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE MATURASE (FRAGMENT).
GN MATR.
OS Pennus boldus.
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Laurales; Monimiacae; Pennus.
OX NCBI_TaxID=63812;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20052162; PubMed=10586879;
RA Qiu Y.-L., Lee J., Bernasconi-Quadrone F., Soltis D.E., Soltis P.S.,
RA Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;
RT "The Earliest Angiosperms: Evidence from Mitochondrial, Plastid and
RT Nuclear Genomes.";
RL Nature 402:404-407(1999).
CC -I- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
CC TRANSCRIPTASE).
CC EMBL; AF197803; AAF14795.1; -
DR InterPro; IPR000442; -
DR InterPro; IPR000477; -
DR Pfam; PF00078; rvt; 1.
DR Pfam; PF01348; Intron_maturas2; 1.
KW Mitochondrion; RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 562 AA; 62910 MW; 735DD7508632B37C CRC64;

Query Match 41.2%; Score 7; DB 8; Length 562;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9
DB 444 ISLHSS 450

RESULT 9
O9t6T0 PRELIMINARY; PRT; 562 AA.
AC O9t6T0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

```

```

DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
GN MATR.
OS Akebia quinata.
OC Mitochondrion.
CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
   Magnoliophyta; eudicotyledons; Ranunculales; Lardizabaleaceae; Akebia.
   NCBI_TaxID=13331;
   [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20052162; PubMed-10586879;
   Qiu Y.-L., Lee J., Bernasconi-Quadroni F., Soltis D.E., Soltis P.S.,
   Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;
   "The Earliest Angiosperms: Evidence from Mitochondrial, Plastid and
   Nuclear Genomes.";
   Nature 402:404-407(1999).
CC -1- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
   TRANSCRIPTASE).
DR EMBL: AF197810; AAF14802.1; -.
DR InterPro: IPR000442; -.
DR InterPro: IPR000477; -.
DR Pfam: PF00078; rvt; 1.
DR Pfam: PF01348; Intron_maturas2; 1.
KW Mitochondrion; RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 562
SQ SEQUENCE 562 AA; 63006 MW; 96758F5324F182BF CRC64;

Query Match          41.2%; Score 7; DB 8; Length 562;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ISLHSS 9
   |||||
DB 447 ISLHSS 453

RESULT 10
ID Q9T6U6 PRELIMINARY; PRT; 565 AA.
AC Q9T6U6;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
DE MATURASE (FRAGMENT).
GN MATR.
OS Trochodendron aralioides.
OC Mitochondrion.
CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
   Magnoliophyta; eudicotyledons; Trochodendraceae; Trochodendroideae;
   OC Trochodendron.
   NCBI_TaxID=4407;
   [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20052162; PubMed-10586879;
   Qiu Y.-L., Lee J., Bernasconi-Quadroni F., Soltis D.E., Soltis P.S.,
   Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;
   "The Earliest Angiosperms: Evidence from Mitochondrial, Plastid and
   Nuclear Genomes.";
   Nature 402:404-407(1999).
CC -1- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
   TRANSCRIPTASE).
DR EMBL: AF197792; AAF14784.1; -.
DR InterPro: IPR000442; -.
DR InterPro: IPR000477; -.
DR Pfam: PF00078; rvt; 2.
DR Pfam: PF01348; Intron_maturas2; 1.
KW Mitochondrion; RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 565
SQ SEQUENCE 565 AA; 63655 MW; FB2E89BE5A5E337C CRC64;

```

```

Query Match          41.2%; Score 7; DB 8; Length 565;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ISLHSS 9
   |||||
DB 426 ISLHSS 432

RESULT 11
ID Q9T701 PRELIMINARY; PRT; 567 AA.
AC Q9T701;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
DE MATURASE (FRAGMENT).
GN MATR.
OS Tholtea tomentosa.
OC Mitochondrion.
CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
   Magnoliophyta; Piperales; Aristolochiaceae; Tholtea.
   NCBI_TaxID=108427;
   [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20052162; PubMed-10586879;
   Qiu Y.-L., Lee J., Bernasconi-Quadroni F., Soltis D.E., Soltis P.S.,
   Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;
   "The Earliest Angiosperms: Evidence from Mitochondrial, Plastid and
   Nuclear Genomes.";
   Nature 402:404-407(1999).
CC -1- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE
   TRANSCRIPTASE).
DR EMBL: AF197733; AAF14725.1; -.
DR InterPro: IPR000442; -.
DR InterPro: IPR000477; -.
DR Pfam: PF00078; rvt; 1.
DR Pfam: PF01348; Intron_maturas2; 1.
KW Mitochondrion; RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 567
SQ SEQUENCE 567 AA; 63427 MW; 6ACEF7DFA547D9AD CRC64;

Query Match          41.2%; Score 7; DB 8; Length 567;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ISLHSS 9
   |||||
DB 444 ISLHSS 450

RESULT 12
ID Q9T6W3 PRELIMINARY; PRT; 577 AA.
AC Q9T6W3;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
DE MATURASE (FRAGMENT).
GN MATR.
OS Litiodendron chilense.
OC Mitochondrion.
CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
   Magnoliophyta; Magnoliales; Magnoliaceae; Litiodendron.
   NCBI_TaxID=3414;
   [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20052162; PubMed-10586879;
   Qiu Y.-L., Lee J., Bernasconi-Quadroni F., Soltis D.E., Soltis P.S.,
   Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;

```

RT "The Earliest Angiosperms: Evidence from Mitochondrial, Plastid and  
 RT Nuclear Genomes.";  
 RL Nature 402:404-407(1999).  
 CC -I- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE  
 CC TRANSCRIPTASE).  
 CC EMBL: AF197774; AAF14766.1; -.  
 DR InterPro: IPR000442; -.  
 DR InterPro: IPR000477; -.  
 DR Pfam: PF00078; rvt; 2.  
 DR Pfam: PF01348; Intron\_maturas2; 1.  
 DR Mitochondrion: RNA-directed DNA polymerase.  
 FT NON\_TER 1  
 FT 577 577  
 SQ SEQUENCE 577 AA: 65029 MW: BEF6A03F2A6FEF43 CRC64;

Query Match 41.2%; Score 7; DB 8; Length 577;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9  
 |||||  
 DB 438 ISLHSS 444

## RESULT 13

ID 09T6V8 PRELIMINARY; PRT; 579 AA.  
 AC 09T6V8;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE MATURASE (FRAGMENT).  
 GN MATR.  
 OS Idiospermum australiense.  
 OG Mitochondrion.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Laurales; Calycanthaceae; Idiospermoideae; Idiospermum.  
 CX NCBT\_TaxID=13573;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20052162; PubMed-10586879;  
 RA Qiu Y.-L., Lee J., Bernasconi-Quadrone F., Soltis D.E., Soltis P.S.,  
 RA Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;  
 RT "The Earliest Angiosperms: Evidence from Mitochondrial, Plastid and  
 RT Nuclear Genomes.";  
 RL Nature 402:404-407(1999).  
 CC -I- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE  
 CC TRANSCRIPTASE).  
 CC EMBL: AF197779; AAF14771.1; -.  
 DR InterPro: IPR000442; -.  
 DR InterPro: IPR000477; -.  
 DR Pfam: PF00078; rvt; 2.  
 DR Pfam: PF01348; Intron\_maturas2; 1.  
 DR Mitochondrion: RNA-directed DNA polymerase.  
 FT NON\_TER 1  
 FT 579 579  
 SQ SEQUENCE 579 AA: 65170 MW: 97CC37AEC3545E82 CRC64;

Query Match 41.2%; Score 7; DB 8; Length 579;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9  
 |||||  
 DB 440 ISLHSS 446

RESULT 14  
 ID 09T6V1 PRELIMINARY; PRT; 580 AA.  
 AC 09T6V1;

DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE MATURASE (FRAGMENT).  
 GN MATR.  
 OS Euptelea polyantha.  
 OG Mitochondrion.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; Ranunculales; Eupteleaceae; Euptelea.  
 CX NCBT\_TaxID=13523;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20052162; PubMed-10586879;  
 RA Qiu Y.-L., Lee J., Bernasconi-Quadrone F., Soltis D.E., Soltis P.S.,  
 RA Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;  
 RT "The Earliest Angiosperms: Evidence from Mitochondrial, Plastid and  
 RT Nuclear Genomes.";  
 RL Nature 402:404-407(1999).  
 CC -I- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE  
 CC TRANSCRIPTASE).  
 CC EMBL: AF197787; AAF14779.1; -.  
 DR InterPro: IPR000442; -.  
 DR InterPro: IPR000477; -.  
 DR Pfam: PF00078; rvt; 2.  
 DR Pfam: PF01348; Intron\_maturas2; 1.  
 DR Mitochondrion: RNA-directed DNA polymerase.  
 FT NON\_TER 1  
 FT 580 580  
 SQ SEQUENCE 580 AA: 65336 MW: AAFCEB5102135416 CRC64;

Query Match 41.2%; Score 7; DB 8; Length 580;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9  
 |||||  
 DB 451 ISLHSS 457

## RESULT 15

ID 09T6X5 PRELIMINARY; PRT; 581 AA.  
 AC 09T6X5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE MATURASE (FRAGMENT).  
 GN MATR.  
 OS Podophyllum peltatum.  
 OG Mitochondrion.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; Ranunculales; Berberidaceae;  
 OC Podophyllum.  
 CX NCBT\_TaxID=35933;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20052162; PubMed-10586879;  
 RA Qiu Y.-L., Lee J., Bernasconi-Quadrone F., Soltis D.E., Soltis P.S.,  
 RA Zanis M., Zimmer E.A., Chen Z., Savolainen V., Chase M.W.;  
 RT "The Earliest Angiosperms: Evidence from Mitochondrial, Plastid and  
 RT Nuclear Genomes.";  
 RL Nature 402:404-407(1999).  
 CC -I- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE  
 CC TRANSCRIPTASE).  
 CC EMBL: AF197762; AAF14754.1; -.  
 DR InterPro: IPR000442; -.  
 DR InterPro: IPR000477; -.  
 DR Pfam: PF00078; rvt; 2.  
 DR Pfam: PF01348; Intron\_maturas2; 1.  
 DR Mitochondrion: RNA-directed DNA polymerase.  
 FT NON\_TER 1  
 FT 581 581

SO SEQUENCE 581 AA: 65395 MM: F3B877598EDED37D CRC64;

Query Match 41.2%; Score 7; DB 8; Length 581;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9  
 |||||  
 Db 445 ISLHSS 451

Search completed: September 15, 2001, 12:50:02  
 Job time: 262 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:45:55 ; Search time 23.18 Seconds  
(without alignments)  
25.123 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_150  
Perfect score: 17  
Sequence: 1 KKISLHSSKEKLRER 17

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	7	41.2	559	1	POTA_MYCGE
2	7	41.2	770	1	KLPA_EMENI
3	6	35.3	68	1	IP3A_YEAST
4	6	35.3	159	1	GRE4_BUCAI
5	6	35.3	160	1	YMT0_YEAST
6	6	35.3	163	1	PGPA_HAEIN
7	6	35.3	177	1	PUR6_PYRHO
8	6	35.3	181	1	NOHB_ECOLI
9	6	35.3	181	1	TERS_LAMBD
10	6	35.3	189	1	NOHA_ECOLI
11	6	35.3	236	1	PLC1_BOVIN
12	6	35.3	236	1	PLC_SHEEP
13	6	35.3	242	1	YTXE_BACSU
14	6	35.3	318	1	Y211_AQUAE
15	6	35.3	328	1	CEBB_CHICK
16	6	35.3	372	1	BIOF_MEPJA
17	6	35.3	389	1	PYRD_DROME
18	6	35.3	397	1	O85A_DROME
19	6	35.3	509	1	G6PD_ANASP
20	6	35.3	509	1	G6PD_NOSUP
21	6	35.3	522	1	RECJ_HELPJ
22	6	35.3	524	1	RECJ_HELPJ
23	6	35.3	526	1	BUTY_BOVIN
24	6	35.3	538	1	NRFA_HAEIN
25	6	35.3	552	1	Y4HP_RHISN
26	6	35.3	561	1	CCB2_DROME
27	6	35.3	571	1	ILV1_BUCAI
28	6	35.3	580	1	YN8X_YEAST
29	6	35.3	598	1	NU5M_BRALA
30	6	35.3	599	1	NU5M_BRALA
31	6	35.3	743	1	PMT5_YEAST
32	6	35.3	795	1	SYFB_ECOLI
33	6	35.3	875	1	UE3A_HUMAN

34	6	35.3	885	1	UE3A_MOUSE
35	6	35.3	917	1	SYI_CAMJE
36	6	35.3	946	1	RHG4_HUMAN
37	6	35.3	1176	1	VPS8_YEAST
38	6	35.3	2210	1	RRPO_TACY
39	6	35.3	4036	1	RRPL_DUGBV
40	6	35.3	4969	1	RYNC_RABIT
41	5	29.4	61	1	DNBI_BFVD
42	5	29.4	62	1	YM45_CAEEL
43	5	29.4	63	1	YORQ_TTVI
44	5	29.4	68	1	RL29_PYRHO
45	5	29.4	72	1	HST1_ECOLI

ALIGNMENTS

RESULT\_1  
POTA\_MYCGE  
ID POTA\_MYCGE STANDARD; PRT; 559 AA.  
AC P47288;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE SPERMIDINE/PUTRESCINE TRANSPORT ATP-BINDING PROTEIN POTA HOMOLOG.  
GN POTA OR MG042.  
OS Mycoplasma genitalium.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 33530 / G-37;  
RX MEDLINE=96026346; PubMed=7569993;  
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
RA Fritchman J.D., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
RA Tomb J.F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,  
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
RT "The minimal gene complement of Mycoplasma genitalium";  
RL Science 270:397-403(1995).  
CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
CC FOR SPERMIDINE/PUTRESCINE. PROBABLY RESPONSIBLE FOR ENERGY  
CC COUPLING TO THE TRANSPORT SYSTEM (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
(ABC TRANSPORTERS).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U39684; AAC71258.1; -  
CC HSSP: P13569; INED.  
CC TIGR: MG042; -  
CC InterPro: IPR001617; -  
CC Pfam: PF00005; ABC\_tran; 1.  
CC PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
CC Transport; ATP-binding; Membrane.  
CC NP\_BIND 40 47 ATP (POTENTIAL).  
SQ SEQUENCE 559 AA; 65192 MW; DB8F04B26A110F6F CRC64;

Query Match 41.2%; Score 7; DB 1; Length 559;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 10 KEKLRR 16

```

Db 173 KEKLRRE 179
|||||
RESULT 2
KLPA_EMENI
ID KLPA_EMENI STANDARD; PRT; 770 AA.
AC P28739;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE KINESIN-LIKE PROTEIN KLPA.
GN KLPA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB20;
RX MEDLINE=93107178; PubMed=8416986;
RA O'Connell M.J., Meluh P.B., Rose M.D., Morris N.R.;
RT "Suppression of the bimC4 mitotic spindle defect by deletion of klpA,
RT a gene encoding a KAR3-related kinesin-like protein in Aspergillus
RT nidulans.";
RL J. Cell Biol. 120:153-162(1993).
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X64603; CAA45887.1; -.
DR PIR; A44337; A44337.
DR HSSP; P17119; 3KAR.
DR InterPro; IPR001752; -.
DR Pfam; PF00225; Kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Microtubules; ATP-binding; Coiled coil.
FT DOMAIN 175 425 COILED COIL (POTENTIAL).
FT DOMAIN 426 770 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).
FT NP_BIND 514 521 ATP (BY SIMILARITY).
SQ SEQUENCE 770 AA; 85800 MW; EFD0F0FF39B5C7EB CRC64;

Query Match 41.2%; Score 7; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KEKLRRE 16
Db 398 KEKLRRE 404
|||||

RESULT 3
IPA3_YEAST
ID IPA3_YEAST STANDARD; PRT; 68 AA.
AC P01094;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PROTEASE A INHIBITOR 3 (PROTEINASE INHIBITOR 1(A)3).
GN PA13 OR YMR174C OR YMR010.04C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=91243884; PubMed=2037077;
RA Schu P., Wolf D.H.;
RT "The proteinase yscA-inhibitor, IA3, gene. Studies of cytoplasmic
RT proteinase inhibitor deficiency on yeast physiology.";
RL FEBS Lett. 283:78-84(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Churher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE.
RA Biedermann K., Montali U., Martin B., Svendsen I., Ottesen M.;
RT "The amino acid sequence of proteinase A inhibitor 3 from baker's
RT yeast.";
RL Carlsberg Res. Commun. 45:225-235(1980).
CC -!- FUNCTION: INHIBITOR FOR YEAST PROTEASE A (YSCA).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60050; CAA42650.1; -.
DR PIR; A01234; IABY3.
DR PIR; S16692; S16692.
DR SGD; S0004786; PA13.
KW Protease inhibitor; Acetylation.
FT MOD_RES 1 1 ACETYLATION.
FT DOMAIN 2 23 NEARLY ALL OF THE INHIBITORY ACTIVITY IS
FT PRESENT IN THIS REGION.
SQ SEQUENCE 68 AA; 7707 MW; 8295BF7AF3E4A01A CRC64;

Query Match 35.3%; Score 6; DB 1; Length 68;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSKEKL 13
Db 14 SSKEKL 19
|||||

RESULT 4
GREA_BUCAI
ID GREA_BUCAI STANDARD; PRT; 159 AA.
AC P57464;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TRANSCRIPTION ELONGATION FACTOR GREA (TRANSCRIPT CLEAVAGE FACTOR
DE GREA).
DE GREA.
GN GREA OR BU384.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TOKYO 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
RL Nature 407:81-86(2000).

```

CC -!- FUNCTION: NECESSARY FOR EFFICIENT RNA POLYMERASE TRANSCRIPTION  
 CC ELONGATION PAST TEMPLATE-ENCODED ARRESTING SITES. THE ARRESTING  
 CC SITES IN DNA HAVE THE PROPERTY OF TRAPPING A CERTAIN FRACTION  
 CC OF ELONGATING RNA POLYMERASES THAT PASS THROUGH, RESULTING IN  
 CC LOCKED TERNARY COMPLEXES. CLEAVAGE OF THE NASCENT TRANSCRIPT BY  
 CC CLEAVAGE FACTORS SUCH AS GRE A OR GRE B ALLOWS THE RESUMPTION OF  
 CC ELONGATION FROM THE NEW 3' TERMINUS. GRE A RELEASES SEQUENCES OF  
 CC 2 TO 3 NUCLEOTIDES (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE GRE A/GRE B FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AP001119; BAB13087.1; -  
 CC PROSITE; PS00829; GREAB\_1; 1.  
 CC PROSITE; PS00830; GREAB\_2; 1.  
 CC KW Transcription regulation; DNA-binding; Coiled coil.  
 CC FT DOMAIN 10 27 COILED COIL (POTENTIAL).  
 CC SQ SEQUENCE 159 AA; 17983 MW; A400D7A95EFD1D3 CRC64;

Query Match 35.3%; Score 6; DB 1; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 EKLRR 16  
 | | | | |  
 DB 13 EKLRR 18

RESULT 5  
 YMT0\_YEAST  
 ID YMT0\_YEAST STANDARD; PRT; 160 AA.  
 AC Q04210;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE HYPOTHETICAL 19.2 KDA PROTEIN IN SUB1-ANGRI INTERGENIC REGION.  
 GN YMR040W OR YMR9532.05.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 CC NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RA Odell C., Bowman S., Barrell B.G., Rajandream M.A.;  
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -!- SIMILARITY: TO YEAST YKL05C.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; Z48502; CAA88406.1; -  
 DR SGD; S0004643; YMR040W.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 3 23 POTENTIAL.  
 FT TRANSMEM 46 66 POTENTIAL.  
 FT TRANSMEM 104 124 POTENTIAL.  
 SQ SEQUENCE 160 AA; 19178 MW; D003EC7412559871 CRC64;

Query Match 35.3%; Score 6; DB 1; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KEKLRR 15  
 | | | | |  
 DB 142 KEKLRR 147

RESULT 6  
 PGPA\_HAEIN  
 ID PGPA\_HAEIN STANDARD; PRT; 163 AA.  
 AC P44157;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE PHOSPHATIDYLGLYCEROPHOSPHATASE A (EC 3.1.3.27).  
 CC PGPA OR H11306  
 CC Haemophilus influenzae.  
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 CC Haemophilus.  
 CC NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerenlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McInney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Haudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 RT influenzae Rd.";  
 RL Science 269:496-512(1995).

CC -!- FUNCTION: SPECIFICALLY HYDROLYZES PHOSPHATIDYLGLYCEROPHOSPHATE (BY  
 CC SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLGLYCEROPHOSPHATE + H(2)O =  
 CC PHOSPHATIDYLGLYCEROL + ORTHOPHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE  
 CC (POTENTIAL).  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U32810; AAC22953.1; -  
 CC TIGR; H11306; -  
 KW Hydrolase; Phospholipid degradation; Transmembrane; Inner membrane.  
 FT TRANSMEM 10 28 POTENTIAL.  
 FT TRANSMEM 35 51 POTENTIAL.  
 FT TRANSMEM 92 116 POTENTIAL.  
 FT TRANSMEM 137 157 POTENTIAL.  
 SQ SEQUENCE 163 AA; 18034 MW; 77CC4D4FC550937B CRC64;

Query Match 35.3%; Score 6; DB 1; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKISLL 6  
 | | | | |  
 DB 8 KKISLL 13

RESULT 7

PUR6\_PYRHO  
ID PUR6\_PYRHO STANDARD; PRT; 177 AA.  
AC O58058;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PHOSPHORIBOSYLAMINOIMIDAZOLE CARBOXYLASE CATALYTIC SUBUNIT  
DE (EC 4.1.1.21) (AIR CARBOXYLASE) (AIRC).  
GN PURE OR PHO320 OR PHCD015.  
OS Pyrococcus horikoshii.  
OC Archaea: Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
OX NCBI\_TaxID=53953;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OT3;  
RX MEDLINE=98344137; PubMed=9679194;  
RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,  
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
RA Masuchi Y., Shizuya H., Kikuchi H.;  
RT "Complete sequence and gene organization of the genome of a hyper-  
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).  
CC -!- FUNCTION: THIS SUBUNIT CAN ALONE TRANSFORM AIR TO CAIR, BUT IN  
CC ASSOCIATION WITH PURK, WHICH POSSESSES AN ATPASE ACTIVITY, AN  
CC ENZYME COMPLEX IS PRODUCED WHICH IS CAPABLE OF CONVERTING AIR TO  
CC CAIR EFFICIENTLY UNDER PHYSIOLOGICAL CONDITION (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: 1-(5-PHOSPHORIBOSYL)-5-AMINO-4-IMIDAZOLE-  
CC CARBOXYLATE = 1-(5-PHOSPHORIBOSYL)-5-AMINOIMIDAZOLE + CO(2).  
CC -!- PATHWAY: SIXTH STEP IN DE NOVO PURINE BIOSYNTHESIS.  
CC -!- SIMILARITY: TO OTHER BACTERIAL PURE, ALSO TO AIR CARBOXYLASE FROM  
CC FUNGI.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AP000002; BAA29394.1; -;  
DR InterPro: IPR000031; -;  
DR Pfam: PF00731; AIRC; 1.  
KW Purine biosynthesis; Lyase; Decarboxylase.  
SQ SEQUENCE 177 AA; 19439 MW; F8648BAA655703F CRC64;  
  
Query Match 35.3%; Score 6; DB 1; Length 177;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 10 KEKLR 15  
Db 145 KEKLR 150  
|||||  
  
RESULT 8  
NOHB\_ECOLI  
ID NOHB\_ECOLI STANDARD; PRT; 181 AA.  
AC P31062; P75721;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE PROPHAGE QSR; DNA PACKAGING PROTEIN NUI HOMOLOG.  
GN NOHB.  
OS Escherichia coli.  
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN [2]  
RP SEQUENCE OF 1-147 FROM N.A.  
RC STRAIN=K12 / W3110;  
RX MEDLINE=92334985; PubMed=1630906;  
RA Kotani H., Kawamura A., Takahashi A., Nakatsuji M., Hiraoka N.,  
RA Nakajima K., Takanami M.;  
RT "Site-specific dissection of E. coli chromosome by lambda terminase.";  
RL Nucleic Acids Res. 20:3357-3360(1992).  
CC -!- SIMILARITY: TO PHAGE LAMBDA DNA PACKAGING PROTEIN NUI.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE000161; AAC73661.1; -;  
DR EMBL; D00928; -; NOT\_ANNOTATED\_CDS.  
DR EcoGene; EGI1635; nohB.  
SQ SEQUENCE 181 AA; 20428 MW; E38D4FBFF174751C CRC64;  
  
Query Match 35.3%; Score 6; DB 1; Length 181;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 11 EKLRE 16  
Db 60 EKLRE 65  
|||||  
  
RESULT 9  
TERS\_LAMB  
ID TERS\_LAMB STANDARD; PRT; 181 AA.  
AC P03707;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE TERMINASE SMALL SUBUNIT (DNA PACKAGING PROTEIN NUI).  
GN NUI.  
OS Bacteriophage lambda.  
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;  
OC Lambda phage group.  
OX NCBI\_TaxID=10710;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83189071; PubMed=6221115;  
RA Sanger F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B.;  
RT "Nucleotide sequence of bacteriophage lambda DNA.";  
RL J. Mol. Biol. 162:729-773(1982).  
RN [2]  
RP ATP-BINDING DOMAIN.  
RX MEDLINE=88172462; PubMed=2965248;  
RA Becker A., Gold M.;  
RT "Prediction of an ATP reactive center in the small subunit, gpNul, of  
RT the phage lambda terminase enzyme.";  
RL J. Mol. Biol. 199:219-222(1988).  
CC -!- FUNCTION: INVOLVED IN THE INITIATION OF LAMBDA DNA PACKAGING INTO  
CC THE PROHEAD. PROCESSES REPLICATING CONCATEMERIC DNA INTO PIECES  
CC OF UNIT LENGTH WITH COHESIVE ENDS.  
CC -!- SUBUNIT: HETEROOLIGOMER OF NUI AND GPA.  
CC -!- SIMILARITY: TO BACTERIOPHAGE 21 TERMINASE SMALL SUBUNIT (GPI).

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J02459; AAA96533.1; -
DR PIR; A04329; JVBPNL.
DR DNA packaging; ATP-binding.
FT NP_BIND 31
SQ SEQUENCE 181 AA; 20441 MW; A25A5EAEF7BF18A4 CRC64;

Query Match 35.3%; Score 6; DB 1; Length 181;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 EKLRR 16
Db 60 EKLRR 65

RESULT 10
NOHA_ECOLI STANDARD; PRT; 189 AA.
ID NOHA_ECOLI
AC P31061; P77152;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROPHAGE QIN DNA PACKAGING PROTEIN NU1 HOMOLOG.
GN NOHA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sempel G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RP SEQUENCE OF 1-147 FROM N.A.
RX STRAIN=K12 / W3110;
RX MEDLINE=92334985; PubMed=1630906;
RA Kotani H., Kawamura A., Takahashi A., Nakatsuji M., Hiraoka N.,
RA Nakajima K., Takanami M.;
RT "Site-specific dissection of E. coli chromosome by lambda terminase.";
RL Nucleic Acids Res. 20:3357-3360(1992).
CC -1- SIMILARITY: TO PHAGE LAMBDA DNA PACKAGING PROTEIN NU1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000252; AAC74621.1; -
DR EMBL; D90798; BAA15252.1; -
DR EMBL; D00927; -; NOT_ANNOTATED_CDS.
DR EcoGene; EG11634; noha.
SQ SEQUENCE 189 AA; 21404 MW; FB5B1F844AB2C7C1 CRC64;

Query Match 35.3%; Score 6; DB 1; Length 189;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 EKLRR 16
Db 60 EKLRR 65

RESULT 11
PLCI_BOVIN STANDARD; PRT; 236 AA.
ID PLCI_BOVIN
AC P09611;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PLACENTAL LACTOGEN I PRECURSOR (BPLP-I).
GN PLI.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90256825; PubMed=2341410;
RX Yamakawa M., Tanaka M., Koyama M., Kagesato Y., Watahiki M.,
RX Yamamoto M., Nakashima K.;
RT "Expression of new members of the prolactin growth hormone gene
RT family in bovine placenta. Isolation and characterization of two
RT prolactin-like cDNA clones.";
RL J. Biol. Chem. 265:8915-8920(1990).
RN [2]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=89207456; PubMed=3242594;
RX Schuler L.A., Shimomura K., Kessler M.A., Zieler C.G., Bremel R.D.;
RT "Bovine placental lactogen: molecular cloning and protein structure.";
RL Biochemistry 27:8443-8448(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91166732; PubMed=2003877;
RX Kessler M.A., Schuler L.A.;
RT "Structure of the bovine placental lactogen gene and alternative
RT splicing of transcripts.";
RL DNA Cell Biol. 10:93-101(1991).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J02840; AAA30709.1; -
DR EMBL; M33268; AAA30739.1; -
DR EMBL; M65218; AAA30612.1; ALT_SEQ.
DR EMBL; M65219; AAA30613.1; ALT_SEQ.

```

DR EMBL; M65224; AAA30611.1; -  
 DR EMBL; M65220; AAA30611.1; JOINED.  
 DR EMBL; M65221; AAA30611.1; JOINED.  
 DR EMBL; M65222; AAA30611.1; JOINED.  
 DR EMBL; M65223; AAA30611.1; JOINED.  
 DR PIR; A36284; A36284.  
 DR HSSP; Q28632; IAN3.  
 DR InterPro; IPR001400; -  
 DR Pfam; PF00103; hormone; 1.  
 DR PRINTS; PR00836; SOMATOTROPIN.  
 DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
 KW Hormone; Placenta; Glycoprotein; Signal.  
 FT SIGNAL 1 36  
 FT CHAIN 37 236 PLACENTAL LACTOGEN I.  
 FT DISULFID 98 214 BY SIMILARITY.  
 FT DISULFID 231 236 BY SIMILARITY.  
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT CONFLICT 37 37 V -> A (IN REF. 2).  
 FT CONFLICT 94 94 V -> F (IN REF. 2 AND 3).  
 FT CONFLICT 170 170 V -> M (IN AAA30613).  
 FT CONFLICT 170 170 V -> M (IN AAA30613).  
 SQ SEQUENCE 236 AA; 26908 MW; 2C46637D8C9D7471 CRC64;  
 Query Match 35.3%; Score 6; DB 1; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3 ISLLHS 8  
 Db 125 ISLLHS 130  
 RESULT 12  
 PLC-SHEEP  
 ID PLC-SHEEP STANDARD; PRT; 236 AA.  
 AC P16038;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE PLACENTAL LACTOGEN PRECURSOR.  
 OS Ovis aries (Sheep).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Caprinae; Ovis.  
 CC NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90114213; PubMed=2608069;  
 RA Colosi P., Thordarson G., Hellmiss R., Singh K., Forsyth I.A.,  
 RA Gluckman P., Wood W.I.;  
 RT "Cloning and expression of ovine placental lactogen.";  
 RL Mol. Endocrinol. 3:1462-1469(1989).  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M31660; AAA31577.1; -  
 DR PIR; A40143; A40143.  
 DR HSSP; Q28632; IAN3.  
 DR InterPro; IPR001400; -  
 DR Pfam; PF00103; hormone; 1.  
 DR PRINTS; PR00836; SOMATOTROPIN.  
 DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
 KW Hormone; Placenta; Signal.

FT SIGNAL 1 36  
 FT CHAIN 37 236 PLACENTAL LACTOGEN.  
 FT DISULFID 97 212 BY SIMILARITY.  
 FT DISULFID 229 234 BY SIMILARITY.  
 SQ SEQUENCE 236 AA; 26695 MW; E6223AFE2FF9BB35 CRC64;  
 Query Match 35.3%; Score 6; DB 1; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3 ISLLHS 8  
 Db 124 ISLLHS 129  
 RESULT 13  
 YTXE\_BACSU  
 ID YTXE\_BACSU STANDARD; PRT; 242 AA.  
 AC P39064;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE HYPOTHETICAL 27.6 KDA PROTEIN IN ACUC 5'REGION (ORFB).  
 GN YTXE.  
 OS Bacillus subtilis.  
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 CC Bacillus/Staphylococcus group; Bacillus.  
 CC NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=95020526; PubMed=7934817;  
 RA Grundy F.J., Waters D.A., Takova T.Y., Henkin T.M.;  
 RT "Identification of genes involved in utilization of acetate and  
 RT acetoin in Bacillus subtilis.";  
 RL Mol. Microbiol. 10:259-271(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98048467; PubMed=9387221;  
 RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;  
 RT "Sequencing and functional annotation of the Bacillus subtilis genes  
 RT in the 200 kb rnb-dnaB region.";  
 RL Microbiology 143:3431-3441(1997).  
 CC -1- FUNCTION: MAY BE INVOLVED IN SOME TRANSPORT FUNCTION.  
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE MOTB FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L17309; AAA68283.1; -  
 DR EMBL; AF008220; AAC00301.1; -  
 DR EMBL; Z99119; CAB14950.1; -  
 DR PIR; S39642; S39642.  
 DR Subtilist; BG10366; YTXE.  
 DR InterPro; IPR001145; -  
 DR Pfam; PF00691; ompa; 1.  
 DR Hypothetical protein; Transport; Transmembrane.  
 KW DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 24 44 POTENTIAL.  
 FT DOMAIN 45 242 EXTRACELLULAR (POTENTIAL).  
 SQ SEQUENCE 242 AA; 27595 MW; 8BA7DDC103C1DF5 CRC64;  
 Query Match 35.3%; Score 6; DB 1; Length 242;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 12 KLRRER 17
DB 2 KLRRER 7

RESULT 14
YZ11_AQAE
ID YZ11_AQAE STANDARD; PRT; 318 AA.
AC O66405;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL PROTEIN AAL1.
GN AAL1.
OS Aquifex aeolicus.
OC Bacteria: Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aufay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus";
RL Nature 392:353-358(1998).
CC -1- SIMILARITY: STRONG TO A.AEOLICUS AA07 AND AA34.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 221646; CAA79760.1; -
DR EMBL; X70813; CAA50144.1; -
DR PIR; S35336; S35336.
DR TRANSFAC; T02022; -
KW Pfam: PF00170; bZIP; 1.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
FT DNA_BIND 260 276
FT DOMAIN 289 317
FT LEUCINE-ZIPPER.
SQ SEQUENCE 328 AA; 35030 MW; 5AAE257F8213671C CRC64;

Query Match 35.3%; Score 6; DB 1; Length 328;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 KLRRER 17
DB 140 KLRRER 145

RESULT 15
CEBB_CHICK
ID CEBB_CHICK STANDARD; PRT; 328 AA.
AC Q05826;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE CCAAT/ENHANCER BINDING PROTEIN BETA (C/EBP BETA) (TRANSCRIPTION
DE FACTOR NF-M) (CCR PROTEIN).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93223673; PubMed=8467792;
RA Katz S., Kowenz-Leutz E., Mueller C., Meese K., Ness S.A.,
RA Leutz A.;
RT "The NF-M transcription factor is related to C/EBP beta and plays a

```

```

RT role in signal transduction, differentiation and leukemogenesis of
RT avian myelomonocytic cells.";
RL EMBO J. 12:1321-1332(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93259145; PubMed=8491193;
RA Burk O., Mink S., Ringwald M., Klempnauer K.H.;
RT "Synergistic activation of the chicken mim-1 gene by v-myb and C/EBP
RT transcription factors.";
RL EMBO J. 12:2027-2038(1993).
CC -1- FUNCTION: HAS A ROLE IN SIGNAL TRANSDUCTION, DIFFERENTIATION AND
CC LEUKEMOGENESIS OF MYELOMONOCYTIC CELLS. BINDS TO THE MGF AND MIM-1
CC PROMOTERS AND ACTIVATES THE TRANSCRIPTION OF THESE GENES.
CC -1- SUBUNIT: BINDS DNA AS A DIMER.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN MYELOMONOCYTIC
CC CELLS.
CC -1- SIMILARITY: TO OTHER BZIP PROTEINS. STRONG, TO OTHER C/EBP
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 221646; CAA79760.1; -
DR EMBL; X70813; CAA50144.1; -
DR PIR; S35336; S35336.
DR TRANSFAC; T02022; -
KW Pfam: PF00170; bZIP; 1.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
FT DNA_BIND 260 276
FT DOMAIN 289 317
FT LEUCINE-ZIPPER.
SQ SEQUENCE 328 AA; 35030 MW; 5AAE257F8213671C CRC64;

Query Match 35.3%; Score 6; DB 1; Length 328;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 KLRRER 17
DB 258 KLRRER 263

Search completed: September 15, 2001, 12:50:38
Job time: 283 sec

```

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:44:30 : Search time 65.58 Seconds  
(without alignments)  
15.715 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_150

Perfect score: 17

Sequence: 1 KKISLHSSKEKLRRER 17

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size: 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

A\_Geneseq\_0601:\*

1: /cgnl\_9/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /cgnl\_9/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /cgnl\_9/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /cgnl\_9/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /cgnl\_9/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /cgnl\_9/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /cgnl\_9/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /cgnl\_9/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
11: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
12: /cgnl\_9/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
13: /cgnl\_9/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /cgnl\_9/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /cgnl\_9/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /cgnl\_9/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /cgnl\_9/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /cgnl\_9/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /cgnl\_9/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /cgnl\_9/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /cgnl\_9/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
22: /cgnl\_9/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	100.0	405	21	AAV79269
2	11	64.7	15	21	AAV79270
3	7	41.2	135	21	AAAG43791
4	7	41.2	250	21	AAAG43790
5	7	41.2	323	21	AAV93937
6	7	41.2	323	22	AAAB6146
7	6	35.3	112	20	AAV30183
8	6	35.3	112	20	AAV30184
9	6	35.3	125	20	AAV12386
10	6	35.3	132	19	AAW79223
11	6	35.3	132	19	AAW79225

12	6	35.3	132	19	AAW56349	Humanised LO-CD2A
13	6	35.3	132	19	AAW56347	LO-CD2A light chain
14	6	35.3	132	20	AAV30186	Sequence of the hu
15	6	35.3	132	20	AAV30181	Amino acid sequenc
16	6	35.3	134	15	AAK59078	LO-CD2A VL and sig
17	6	35.3	151	20	AAV35951	Extended human sec
18	6	35.3	151	20	AAV36094	Extended human sec
19	6	35.3	152	21	AAV54945	Arabidopsis thalia
20	6	35.3	250	21	AAAB42198	Human ORF ORF1962
21	6	35.3	253	15	AAK59083	Protein encoded fr
22	6	35.3	280	21	AAAB25192	Euclalyptus grandis
23	6	35.3	280	22	AAAB5671	Novel protein kina
24	6	35.3	284	21	AAAG22161	Arabidopsis thalia
25	6	35.3	291	21	AAAG18003	Arabidopsis thalia
26	6	35.3	292	19	AAAG49791	Arabidopsis thalia
27	6	35.3	292	19	AAW54348	Arabidopsis thalia
28	6	35.3	305	21	AAAB56616	Streptococcal sal
29	6	35.3	311	21	AAAG49790	Human prostate can
30	6	35.3	320	21	AAAG18002	Arabidopsis thalia
31	6	35.3	329	21	AAAG22160	Arabidopsis thalia
32	6	35.3	333	21	AAAG22159	Arabidopsis thalia
33	6	35.3	337	22	AAAB60661	Arabidopsis thalia
34	6	35.3	348	21	AAAG18001	Human apoptosis-re
35	6	35.3	360	21	AAAG49789	Arabidopsis thalia
36	6	35.3	397	21	AAAB26437	Arabidopsis thalia
37	6	35.3	397	21	AAAB20934	Drosophila melanog
38	6	35.3	439	21	AAAB4905	Drosophila odorant
39	6	35.3	524	20	AAW97815	A human proliferat
40	6	35.3	526	20	AAW97812	Guinea pig butyrop
41	6	35.3	571	21	AAAG13369	Bovine butyrophili
42	6	35.3	572	21	AAAG13368	Arabidopsis thalia
43	6	35.3	591	21	AAV94911	Arabidopsis thalia
44	6	35.3	591	22	AAAB60459	Human secreted pro
45	6	35.3	762	21	AAAG52383	Human cell cycle a
						Arabidopsis thalia

#### ALIGNMENTS

RESULT 1		
ID	AAV79269	standard; Protein; 405 AA.
XX		
AC	AAV79269;	
XX		
DT	03-JUL-2000	(first entry)
XX		
DE	Human testis-specific transcription factor PHELI.	
XX		
KW	PHELI: human; testis-specific; transcription factor;	
KW	prostate cancer; bladder cancer; ovary cancer; testicular cancer;	
KW	therapy; diagnosis; vaccine.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	Peptide	134..150
FT		/note="nuclear localization signal"
FT	Peptide	163..169
FT		/note="nuclear localization signal"
FT	Domain	140..189
FT		/note="basic Helix-Loop-Helix domain"
PN	WO200012709-A2.	
XX		
ED	09-MAR-2000.	
XX		
PF	31-AUG-1999;	99WO-US20137.
XX		
PR	31-AUG-1998;	98US-0098610.
PR	31-OCT-1998;	98US-0106524.
XX		
PA	(UROG-) UROGENESYS INC.	

PA (AFAR/) AFAR D E.  
PA (HUBE/) HUBERT R S.  
PA (RAIT/) RAITANO A B.  
XX  
PI Afar DE, Hubert RS, Raitano AB;  
XX  
DR WPI; 2000-237872/20.  
DR N-PSDB; AA294275.  
XX  
PT Testis specific Helix loop Helix proteins expressed in cancers and  
PT useful for the prevention, diagnosis and treatment of prostate, bladder  
PT and ovarian tumors -  
XX  
PS Claim 1; Fig 2A-D; 62pp; English.  
XX  
CC This sequence is that of human PHELIIX, a novel basic Helix loop  
CC Helix protein thought to act as a transcription factor. PHELIIX  
CC normally exhibits a testis-specific expression pattern but is  
CC up-regulated in prostate and other types of cancer. The invention  
CC provides diagnostic and therapeutic methods useful in the  
CC management of various cancers which express PHELIIX, including  
CC prostate cancer, bladder cancer, ovarian cancer and testicular  
CC cancer, including therapies aimed at inhibition the transcription,  
CC translation, processing or function of PHELIIX. The expression  
CC pattern of PHELIIX suggests that is an ideal target for a cancer  
CC vaccine approach to prostate cancer. PHELIIX protein can also be  
CC used to screen for agonists and antagonists of therapeutic value  
CC and to raise antibodies.  
XX  
SQ Sequence 405 AA;

Query Match 100.0%; Score 17; DB 21; Length 405;  
Best Local Similarity 100.0%; Pred. No. 6.5e-09;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKSILHSSKEKLRER 17  
Db 134 KKSILHSSKEKLRER 150

RESULT 2  
AAV79270  
ID AAV79270 standard; Peptide; 15 AA.  
XX  
AC AAV79270;  
XX  
DT 03-JUL-2000 (first entry)  
XX  
DE PHELIIX peptide used to raise antibody.  
XX  
KW PHELIIX; human; testis-specific; transcription factor;  
KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
KW therapy; diagnosis; vaccine; antibody.  
XX  
OS Homo sapiens.  
XX  
PN WO200012709-A2.  
PD  
XX  
PD 09-MAR-2000.  
XX  
PF 31-AUG-1999; 99WO-US20137.  
XX  
PR 31-AUG-1998; 98US-0098610.  
PR 31-OCT-1998; 98US-0106524.  
XX  
PA (UROG-) UROGENESTIS INC.  
PA (AFAR/) AFAR D E.  
PA (HUBE/) HUBERT R S.  
PA (RAIT/) RAITANO A B.  
XX  
PI Afar DE, Hubert RS, Raitano AB;  
XX

DR WPI; 2000-237872/20.  
XX  
XX Testis specific Helix loop Helix proteins expressed in cancers and  
PT useful for the prevention, diagnosis and treatment of prostate, bladder  
PT and ovarian tumors -  
XX  
PS Example 5; Page 36; 62pp; English.  
XX  
CC The present sequence is that of a peptide derived from human  
CC PHELIIX (see AAV79269), a novel transcription factor that is  
CC normally expressed only in testis tissue, but which is up-regulated  
CC in prostate and some other cancers. The peptide was conjugated to  
CC keyhole limpet haemocyanin and used to raise polyclonal antiserum  
CC in rabbit. The antiserum demonstrated specificity for PHELIIX and  
CC may therefore be useful for assessing the expression of PHELIIX in  
CC patient samples.  
XX  
SQ Sequence 15 AA;

Query Match 64.7%; Score 11; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 HSSKEKLRER 17  
Db 1 hsskekrlrrer 11

RESULT 3  
AAG43791  
ID AAG43791 standard; Protein; 135 AA.  
XX  
AC AAG43791;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54775.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-01343768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0138119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.

```
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 41.2%; Score 7; DB 21; Length 135;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SSKEXLR 14
Db 129 sskExlr 135

RESULT 4
AAG43790
ID AAG43790 standard; Protein; 250 AA.
XX
AC AAG43790;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54774.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
```

```
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
```

PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match 41.2%; Score 7; DB 21; Length 250;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 8 SSKERL3 14  
Db 244 ssketr 250  
  
RESULT 5  
AAV9397 standard; Protein: 323 AA.  
ID AAV9397  
XX AAV9397;  
AC AAV9397;  
XX 08-AUG-2000 (first entry)  
DT 08-AUG-2000 (first entry)  
XX 08-AUG-2000 (first entry)  
DE Human PRO1298 (UNQ666) amino acid sequence SEQ ID NO:210.  
XX Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;  
KW Transmembrane; secretion; Immunoadhesion; pharmaceutical; screening.  
XX Homo sapiens.  
OS  
PN WO200012708-A2.  
PD 09-MAR-2000.  
XX 09-MAR-2000.  
XX 01-SEP-1999; 99WO-US20111.  
PF 01-SEP-1999; 99US-0098716.  
XX 01-SEP-1999; 99US-0098716.  
PR 01-SEP-1999; 99US-0098750.  
PR 02-SEP-1999; 99US-0098803.  
PR 02-SEP-1999; 99US-0098821.  
PR 02-SEP-1999; 99US-0098843.  
PR 09-SEP-1999; 99US-0099536.  
PR 09-SEP-1999; 99US-0099596.  
PR 09-SEP-1999; 99US-0099598.  
PR 09-SEP-1999; 99US-0099642.  
PR 09-SEP-1999; 99US-0099741.  
PR 10-SEP-1999; 99US-0099754.  
PR 10-SEP-1999; 99US-0099763.  
PR 10-SEP-1999; 99US-0099792.  
PR 10-SEP-1999; 99US-0099808.  
PR 10-SEP-1999; 99US-0099812.  
PR 10-SEP-1999; 99US-0099815.  
PR 10-SEP-1999; 99US-0099816.  
PR 15-SEP-1999; 99US-0100385.  
PR 15-SEP-1999; 99US-0100388.  
PR 15-SEP-1999; 99US-0100390.  
PR 16-SEP-1999; 99US-0100584.  
PR 16-SEP-1999; 99US-0100627.  
PR 16-SEP-1999; 99US-0100661.  
PR 16-SEP-1999; 99US-0100662.  
PR 16-SEP-1999; 99US-0100664.  
PR 16-SEP-1999; 99US-0100683.  
PR 17-SEP-1999; 99US-0100710.  
PR 17-SEP-1999; 99US-0100711.  
PR 17-SEP-1999; 99US-0100719.  
PR 17-SEP-1999; 99US-0100919.  
PR 17-SEP-1999; 99US-0100930.  
PR 18-SEP-1999; 99US-0100848.  
PR 18-SEP-1999; 99US-0100849.  
PR 18-SEP-1999; 99US-0101014.  
PR 18-SEP-1999; 99US-0101068.

PR 18-SEP-1998; 98US-0101071.  
 PR 22-SEP-1998; 98US-0101279.  
 PR 23-SEP-1998; 98US-0101471.  
 PR 23-SEP-1998; 98US-0101472.  
 PR 23-SEP-1998; 98US-0101474.  
 PR 23-SEP-1998; 98US-0101475.  
 PR 23-SEP-1998; 98US-0101476.  
 PR 23-SEP-1998; 98US-0101477.  
 PR 23-SEP-1998; 98US-0101478.  
 PR 24-SEP-1998; 98US-0101738.  
 PR 24-SEP-1998; 98US-0101741.  
 PR 24-SEP-1998; 98US-0101743.  
 PR 24-SEP-1998; 98US-0101915.  
 PR 24-SEP-1998; 98US-0101916.  
 PR 29-SEP-1998; 98US-0102207.  
 PR 29-SEP-1998; 98US-0102240.  
 PR 29-SEP-1998; 98US-0102307.  
 PR 29-SEP-1998; 98US-0102330.  
 PR 29-SEP-1998; 98US-0102331.  
 PR 30-SEP-1998; 98US-0102484.  
 PR 30-SEP-1998; 98US-0102487.  
 PR 30-SEP-1998; 98US-0102570.  
 PR 30-SEP-1998; 98US-0102571.  
 PR 01-OCT-1998; 98US-0102684.  
 PR 01-OCT-1998; 98US-0102687.  
 PR 02-OCT-1998; 98US-0102965.  
 PR 06-OCT-1998; 98US-0103258.  
 PR 07-OCT-1998; 98US-0103449.  
 PR 07-OCT-1998; 98US-0103314.  
 PR 07-OCT-1998; 98US-0103315.  
 PR 07-OCT-1998; 98US-0103328.  
 PR 07-OCT-1998; 98US-0103395.  
 PR 07-OCT-1998; 98US-0103396.  
 PR 07-OCT-1998; 98US-0103401.  
 PR 08-OCT-1998; 98US-0103633.  
 PR 08-OCT-1998; 98US-0103678.  
 PR 08-OCT-1998; 98US-0103679.  
 PR 08-OCT-1998; 98US-0103711.  
 PR 14-OCT-1998; 98US-0104257.  
 PR 20-OCT-1998; 98US-0104987.  
 PR 20-OCT-1998; 98US-0105000.  
 PR 20-OCT-1998; 98US-0105002.  
 PR 21-OCT-1998; 98US-0105104.  
 PR 22-OCT-1998; 98US-0105169.  
 PR 22-OCT-1998; 98US-0105266.  
 PR 26-OCT-1998; 98US-0105693.  
 PR 26-OCT-1998; 98US-0105694.  
 PR 27-OCT-1998; 98US-0105807.  
 PR 27-OCT-1998; 98US-0105881.  
 PR 27-OCT-1998; 98US-0105882.  
 PR 27-OCT-1998; 98US-0106062.  
 PR 28-OCT-1998; 98US-0106023.  
 PR 28-OCT-1998; 98US-0106029.  
 PR 28-OCT-1998; 98US-0106030.  
 PR 28-OCT-1998; 98US-0106032.  
 PR 28-OCT-1998; 98US-0106033.  
 PR 28-OCT-1998; 98US-0106178.  
 PR 29-OCT-1998; 98US-0106248.  
 PR 29-OCT-1998; 98US-0106384.  
 PR 29-OCT-1998; 98US-0108500.  
 PR 30-OCT-1998; 98US-0106464.  
 PR 03-NOV-1998; 98US-0106856.  
 PR 03-NOV-1998; 98US-0106902.  
 PR 03-NOV-1998; 98US-0106905.  
 PR 03-NOV-1998; 98US-0106919.  
 PR 03-NOV-1998; 98US-0106932.  
 PR 03-NOV-1998; 98US-0106934.  
 PR 10-NOV-1998; 98US-0107783.  
 PR 17-NOV-1998; 98US-0108775.  
 PR 17-NOV-1998; 98US-0108779.  
 PR 17-NOV-1998; 98US-0108787.  
 PR 17-NOV-1998; 98US-0108788.  
 PR 17-NOV-1998; 98US-0108801.

PR 17-NOV-1998; 98US-0108802.  
 PR 17-NOV-1998; 98US-0108806.  
 PR 17-NOV-1998; 98US-0108807.  
 PR 17-NOV-1998; 98US-0108867.  
 PR 17-NOV-1998; 98US-0108869.  
 PR 18-NOV-1998; 98US-0108925.  
 PR 18-NOV-1998; 98US-0108848.  
 PR 18-NOV-1998; 98US-0108849.  
 PR 18-NOV-1998; 98US-0108850.  
 PR 18-NOV-1998; 98US-0108851.  
 PR 18-NOV-1998; 98US-0108852.  
 PR 18-NOV-1998; 98US-0108858.  
 PR 18-NOV-1998; 98US-0108904.  
 XX (GETH ) GENENTECH INC.  
 PA Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;  
 XX WPT, 2000-237871/20.  
 PI N-PSDB; AAA37079.  
 DR  
 DR  
 XX  
 PT New mammalian DNA sequences encoding transmembrane, receptor or  
 PT secreted PRO polypeptides, useful for screening of potential peptide or  
 PT small molecule inhibitors of the relevant receptor/ligand interactions  
 XX  
 PS Claim 12; Fig 116; 773pp; English.  
 XX  
 CC AAA37022 to AAA37144 encode the new isolated human transmembrane,  
 CC receptor or secreted PRO polypeptides given in AA99340 to AA99462. The  
 CC transmembrane and receptor PRO proteins can be used for screening of  
 CC potential peptide or small molecule inhibitors of the relevant  
 CC receptor/ligand interactions. The polypeptides and nucleotide sequences  
 CC encoding then have various industrial applications, including uses as  
 CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent  
 CC PCR primers and hybridisation probes used in the isolation of the PRO  
 CC polypeptides from the present invention.  
 XX  
 SQ Sequence 323 AA:  
 Query Match 41.2%; Score 7; DB 21; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 KISLHLS 8  
 Db 215 KISLHLS 221  
 RESULT 6  
 AAB66146  
 ID AAB66146 standard; protein; 323 AA.  
 AC AAB66146;  
 XX  
 DT 02-APR-2001 (first entry)  
 XX  
 DE Protein of the invention #58.  
 XX  
 KW Secreted; transmembrane; gene therapy.  
 XX  
 CS Unidentified.  
 XX  
 PN WO200078961-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 18-FEB-2000; 2000WO-US04342.  
 XX  
 PR 23-JUN-1999; 99US-0141037.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 01-SEP-1999; 99WO-US20111.  
 PR 29-OCT-1999; 99US-0162506.

PR 30-NOV-1999; 99WO-US28313.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 05-JAN-2000; 2000WO-US00219.  
 PR 06-JAN-2000; 2000WO-US00376.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S,  
 PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ,  
 PI Pan J, Peoni NF, Roy MA, Smith V, Stewart TA, Tumas D,  
 PI Watanabe CK, Williams PM, Wood WI;  
 XX  
 DR WPI: 2001-071395/08.  
 XX  
 PT Secreted and transmembrane proteins and nucleic acids designated PRO,  
 PT useful as hybridization probes, in chromosome and gene mapping and gene  
 PT therapy -  
 XX  
 PS Claim 1; Fig 116; 787pp; English.  
 XX  
 CC The present invention relates to secreted and transmembrane proteins.  
 CC These proteins and the DNA encoding them may be used as hybridization  
 CC probes, in chromosome and gene mapping and in the generation of  
 CC anti-sense RNA and DNA. They may also be used to generate either  
 CC transgenic animals or knockout animals which are in turn useful for  
 CC development and screening of therapeutically useful reagents.  
 CC The nucleic acids may also be used in gene therapy.  
 XX  
 SQ Sequence 323 AA;

Query Match 41.2%; Score 7; DB 22; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 KSLHS 8  
 Db 215 KSLHS 221

RESULT 7  
 ID AAY30183 standard; Protein; 112 AA.  
 XX  
 AC AAY30183;  
 XX  
 DT 29-OCT-1999 (first entry)  
 XX  
 DE Amino acid sequence of light chain variable region of LO-CD2a.  
 XX  
 KW Antibody LO-CD2a; CD2 antigen; T-lymphocyte; humanised antibody;  
 KW T-cell-mediated immune response; graft rejection; autoimmune disease;  
 KW graft-versus-host disease; T cell; natural killer cell.  
 XX  
 OS Rattus sp.  
 XX  
 PN US5951983-A.  
 XX  
 PD 14-SEP-1999.  
 XX  
 PE 07-JUN-1995; 95US-0477989.  
 XX  
 PR 07-JUN-1995; 95US-0477989.  
 PR 05-MAR-1993; 93US-0027008.  
 PR 09-SEP-1993; 93US-0119032.  
 PR 29-MAR-1995; 95US-0407009.  
 XX  
 PA (BIOT-) BIO TRANSPLANT INC.  
 PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.  
 XX  
 PI Bazin H, Kaplan R, Kieber-Emmons T, Latine D, Postema CE;  
 PI White-Scharf ME;

XX  
 DR WPI: 1999-526991/44.  
 XX  
 PT Antibody mediated inhibition of T cell immune response  
 XX  
 PS Disclosure; Column 87; 104pp; English.  
 XX  
 CC The present sequence represents the light chain variable region of rat  
 CC monoclonal antibody LO-CD2a. LO-CD2a binds to an epitope of a CD2  
 CC antigen T-lymphocytes. The specification describes humanised LO-CD2a  
 CC antibodies. The humanized LO-CD2a antibody comprises the human constant  
 CC regions, a light chain framework region derived from a human antibody,  
 CC a heavy chain framework region derived from a human antibody, heavy and  
 CC light chain complementarity determining regions (CDRs) of the non-human  
 CC monoclonal antibody produced by the cell line deposited as ATCC HB11423.  
 CC The humanised antibodies are used in a method for treating a patient to  
 CC inhibit a T-cell-mediated immune response. The method is useful for the  
 CC treatment or prevention of graft rejection and graft-versus-host disease,  
 CC as well as in the treatment of autoimmune diseases which are mediated  
 CC by the activation and proliferation of T cells or natural killer cells.  
 XX  
 SQ Sequence 112 AA;

Query Match 35.3%; Score 6; DB 20; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 4 SLHS 9  
 Db 28 SLHS 33

RESULT 8  
 ID AAY30184 standard; Protein; 112 AA.  
 XX  
 AC AAY30184;  
 XX  
 DT 29-OCT-1999 (first entry)  
 XX  
 DE Amino acid sequence of humanised light chain variable region of LO-CD2a.  
 XX  
 KW Antibody LO-CD2a; CD2 antigen; T-lymphocyte; humanised antibody;  
 KW T-cell-mediated immune response; graft rejection; autoimmune disease;  
 KW graft-versus-host disease; T cell; natural killer cell.  
 XX  
 OS Synthetic.  
 OS Rattus sp.  
 OS Homo sapiens.  
 XX  
 PN US5951983-A.  
 XX  
 PD 14-SEP-1999.  
 XX  
 PE 07-JUN-1995; 95US-0477989.  
 XX  
 PR 07-JUN-1995; 95US-0477989.  
 PR 05-MAR-1993; 93US-0027008.  
 PR 09-SEP-1993; 93US-0119032.  
 PR 29-MAR-1995; 95US-0407009.  
 XX  
 PA (BIOT-) BIO TRANSPLANT INC.  
 PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.  
 XX  
 PI Bazin H, Kaplan R, Kieber-Emmons T, Latine D, Postema CE;  
 PI White-Scharf ME;  
 DR WPI: 1999-526991/44.  
 XX  
 PT Antibody mediated inhibition of T cell immune response  
 XX  
 PS Claim 2; Column 87-89; 104pp; English.

XX The present sequence represents the light chain variable region of a  
CC humanised rat monoclonal antibody LO-CD2a. LO-CD2a binds to an epitope of  
CC a CD2 antigen T-lymphocytes. The specification describes humanised  
CC LO-CD2a antibodies. The humanized LO-CD2a antibody comprises the human  
CC constant regions, a light chain framework region derived from a human  
CC antibody, a heavy chain framework region derived from a human antibody,  
CC heavy and light chain complementarity determining regions (CDRs) of the  
CC non-human monoclonal antibody produced by the cell line deposited as  
CC ATCC HB11423. The humanised antibodies are used in a method for treating  
CC a patient to inhibit a T-cell-mediated immune response. The method is  
CC useful for the treatment or prevention of graft rejection and  
CC diseases which are mediated by the activation and proliferation of  
CC T cells or natural killer cells.

XX Sequence 112 AA:

QY 4 SLHSS 9  
DB 28 slhss 33

Query Match 35.3%; Score 6; DB 20; Length 112;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 9  
AAV12386  
ID AAV12386 standard; Protein; 125 AA.

XX AAV12386;  
XX 17-JUN-1999 (first entry)

DE Human 5' EST secreted protein SEQ ID NO:417.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;  
KW forensic; gene therapy; chromosome mapping; signal peptide;  
KW upstream regulatory sequence; cytokine activity; cell proliferation;  
KW differentiation; haematopoiesis regulation; tissue growth regulation;  
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;  
KW thrombolytic; anti-inflammatory; tumour inhibition.

XX Homo sapiens.  
OS  
XX MO9906548-A2.  
XX 11-FEB-1999.  
XX 31-JUL-1998; 98WO-1B01222.  
XX 01-AUG-1997; 97US-0905135.  
XX (GEST ) GENSET.  
XX Ducleert A, Dumas Milne Edwards J, Lacroix B;  
XX WPI: 1999-153778/13.  
XX N-PSDB; AAX41219.

XX New nucleic acids encoding human secreted proteins - obtained from  
PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,  
PT kidney, lung, umbilical cord, placenta and colon tissue

XX Claim 27; Page 736; 824pp; English.

XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for  
CC human secreted proteins, and encode the proteins given in AAV12261 to  
CC AAV12514, respectively. The proteins given represent the signal peptide  
CC and an N-terminal fragment of a secreted protein. The nucleic acid  
CC sequences can be used for producing secreted human gene products. They

CC can also be used to develop products for diagnosis and therapy. The  
CC proteins obtained may have cytokine activity, cell  
CC proliferation/differentiation activity, haematopoiesis regulating  
CC activity, tissue growth regulating activity, reproductive hormone  
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and  
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory  
CC activity, tumour inhibition activity or other activities. The products  
CC can be used in forensic, gene therapy and chromosome mapping procedures.  
CC The sequences can also be used for obtaining corresponding promoter  
CC sequences. The nucleic acids encoding the signal peptide can be used for  
CC directing extracellular secretion of a polypeptide or the insertion of a  
CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 125 AA:

QY 5 LHSK 10  
DB 48 lhsak 53

Query Match 35.3%; Score 6; DB 20; Length 125;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10  
AAW79223  
ID AAW79223 standard; Protein; 132 AA.

XX AAW79223;  
XX 21-DEC-1998 (first entry)

DE Amino acid sequence of chimeric LO-CD2a antibody VL chain.

XX Monoclonal antibody; Mab; LO-CD2a; humanised antibody; chimeric;  
KW T-cell; immune response; CD2 antigen; graft-versus-host disease;  
KW human lymphocyte; transplant rejection; autoimmune disease.

XX Rattus sp.  
OS  
XX Key  
FH Location/Qualifiers  
FT 1..20  
FT /note= "leader sequence"  
FT Region  
FT 1..23  
FT /note= "Framework 1"  
FT Region  
FT 24..39  
FT /note= "complementarity determining region (CDR) 1"  
FT Region  
FT 40..54  
FT /note= "Framework 2"  
FT Region  
FT 55..61  
FT /note= "complementarity determining region (CDR) 2"  
FT Region  
FT 62..93  
FT /note= "Framework 3"  
FT Region  
FT 94..102  
FT /note= "complementarity determining region (CDR) 3"  
FT Region  
FT 103..112  
FT /note= "Framework 4"

XX US617311-A.  
XX 06-OCT-1998.  
XX 07-JUN-1995; 95US-0472281.  
XX 07-JUN-1995; 95US-0472281.  
XX 07-JUN-1995; 95US-0472281.  
XX 05-MAR-1993; 93US-0027008.  
XX 09-SEP-1993; 93US-0119032.  
XX 29-MAR-1995; 95US-0407009.

XX (UWLO-) UNITIV CATHOLIQUE LOUVAIN.  
XX Bazin H, Latime D;





PI Bazin H, Kaplan R, Kieber-Emmons T, Latime D, Postema CE;  
 PI White-Schaf ME;  
 XX WPI: 1998-168898/15.  
 DR N-PSDB: AAV22853.  
 XX  
 XX Humanised antibody - comprises complementarity determining region  
 PT from LO-CD2a, useful to prevent or inhibit graft versus host or  
 PT auto-immune disease  
 XX  
 PS Claim 2: Fig 32A-C; 133pp; English.  
 XX  
 CC This polypeptide comprises the light chain variable region of  
 CC humanised LO-CD2a. LO-CD2a is a rat anti-CD2 monoclonal  
 CC antibody produced by the hybridoma cell line ATCC HB 11423. The  
 CC humanised VL has human HUM5400 framework regions and rat LO-CDR2a  
 CC complementarity determining regions. Amino acids 9, 12, 41, 42,  
 CC 50, 51 and 82 of the rat LO-CD2a VL framework are retained to  
 CC maintain binding specificity. The DNA sequence (see AAV22853)  
 CC encoding the humanised VL was constructed from overlapping  
 CC oligonucleotides and has been inserted into the MRC light chain  
 CC vector. The VH region (see AAW56350) was similarly obtained.  
 CC Humanised LO-CD2a can be expressed in mammalian host cells. It can  
 CC be used to inhibit an immune response, specifically human T cell  
 CC activation and proliferation, in a patient, and to inhibit  
 CC rejection of a graft in a patient (claimed), useful for preventing  
 CC or inhibiting graft versus host or autoimmune disease. Chimeric  
 CC LO-CD2a antibody (see AAW56347-48) is also claimed.  
 CC  
 XX Sequence 132 AA:  
 SQ  
 Query Match 35.3%; Score 6; DB 19; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 SLHSS 9  
 |||||  
 Db 48 slhss 53

RESULT 13  
 AAW56347  
 ID AAW56347 standard; Protein; 132 AA.  
 XX  
 AC AAW56347;  
 XX  
 DF 17-AUG-1998 (first entry)  
 XX  
 DE LO-CD2a light chain variable region.  
 XX  
 KW LO-CD2a: monoclonal antibody; CD2; rat; chimeric antibody;  
 KW humanised antibody; antibody engineering; graft rejection;  
 KW graft versus host disease; autoimmune disease; therapy.  
 XX  
 OS Rattus sp.  
 XX  
 XX Location/Qualifiers  
 FH 1..20  
 FT /Label= Sig\_peptide  
 FT 21..43  
 FT /note= "framework region 1"  
 FT 44..59  
 FT /note= "complementarity determining region 1"  
 FT 60..74  
 FT /note= "framework region 2"  
 FT 75..81  
 FT /note= "complementarity determining region 2"  
 FT 82..113  
 FT /note= "framework region 3"  
 FT 114..122  
 FT /note= "complementarity determining region 3"  
 FT 123..132  
 FT Region

FT /note= "framework region 4"  
 FT Misc-difference 105  
 FT /note= "encoded by GYT"  
 FT  
 PN WO9807444-A1.  
 XX  
 XX 26-FEB-1998.  
 XX  
 XX 16-AUG-1996; 96WO-US13281.  
 XX  
 XX 16-AUG-1996; 96WO-US13281.  
 XX  
 XX 16-AUG-1996; 96WO-US13281.  
 XX  
 XX (BIOT-) BIOTRANSPLANT INC.  
 XX (UYLO-) UNIV CATHOLIQUE LOUVAIN.  
 XX  
 PI Bazin H, Kaplan R, Kieber-Emmons T, Latime D, Postema CE;  
 PI White-Schaf ME;  
 XX WPI: 1998-168898/15.  
 DR N-PSDB: AAV22851.  
 XX  
 XX Humanised antibody - comprises complementarity determining region  
 PT from LO-CD2a, useful to prevent or inhibit graft versus host or  
 PT auto-immune disease  
 XX  
 PS Claim 8: Fig 29A-C; 133pp; English.  
 XX  
 CC This polypeptide comprises the light chain variable region of  
 CC LO-CD2a, a rat anti-CD2 monoclonal antibody produced by the  
 CC hybridoma cell line deposited as ATCC HB 11423. The VL gene  
 CC (see AAV22851) was obtained by PCR amplification of LO-CD2a cDNA  
 CC (see AAV22855-56), and has been inserted into the light chain vector  
 CC hcwv-vllys-kr-neo, containing the genomic clone of the human kappa  
 CC constant region. A novel chimeric antibody comprises a human  
 CC constant region and the variable regions (see also AAW56348) of rat  
 CC LO-CD2a. The chimeric antibody binds the same epitope or portion  
 CC on human lymphocytes as LO-CD2a. It can be used to inhibit an  
 CC immune response, specifically human T cell activation and  
 CC proliferation, in a patient, and to inhibit rejection of a graft in  
 CC a patient (claimed), useful for preventing or inhibiting graft  
 CC versus host or autoimmune disease. Humanised LO-CD2a antibody (see  
 CC AAW56349-50) is also claimed.  
 CC  
 XX Sequence 132 AA:  
 SQ  
 Query Match 35.3%; Score 6; DB 19; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 SLHSS 9  
 |||||  
 Db 48 slhss 53

RESULT 14  
 AAY30186  
 ID AAY30186 standard; Protein; 132 AA.  
 XX  
 AC AAY30186;  
 XX  
 DF 15-NOV-1999 (first entry)  
 XX  
 DE Sequence of the humanised LO-CD2a light chain variable region.  
 XX  
 KW Antibody LO-CD2a; CD2 antigen; T-lymphocyte; humanised antibody;  
 KW T cell-mediated immune response; graft rejection; autoimmune disease;  
 KW graft-versus-host disease; T cell; natural killer cell.  
 XX  
 OS Synthetic.  
 OS Rattus sp.  
 OS Homo sapiens.  
 XX

```

PN  US5951983-A.
XX
XX  14-SEP-1999.
XX
XX
XX  07-JUN-1995; 95US-0477989.
XX
XX  07-JUN-1995; 95US-0477989.
PR  05-MAR-1993; 93US-0027008.
PR  09-SEP-1993; 93US-0119032.
PR  29-MAR-1995; 95US-0407009.
XX
XX  (BIOT-) BIO TRANSPLANT INC.
PA  (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX
PI  Bazin H, Kaplan R, Kieber-Emmons T, Latime D, Postema CE;
PI  White-Scharf ME;
DR  MPI: 1999-526991/44.
DR  N-PSDB: AA210196.
XX
XX  Antibody mediated inhibition of T cell immune response
XX
XX  Example 7: Column 91-92; 104pp; English.
XX
XX  The present sequence represents humanised LO-CD2a light chain variable
CC  region of rat monoclonal antibody LO-CD2a. LO-CD2a binds to an epitope of
CC  a CD2 antigen T-lymphocytes. The specification describes humanised
CC  LO-CD2a antibodies. The humanized LO-CD2a antibody comprises the human
CC  constant regions, a light chain framework region derived from a human
CC  antibody, a heavy chain framework region derived from a human antibody,
CC  heavy and light chain complementarity determining regions (CDRs) of the
CC  non-human monoclonal antibody produced by the cell line deposited as
CC  ATCC HB11423. The humanised antibodies are used in a method for treating
CC  a patient to inhibit a T-cell-mediated immune response. The method is
CC  useful for the treatment or prevention of graft rejection and
CC  graft-versus-host disease, as well as in the treatment of autoimmune
CC  diseases which are mediated by the activation and proliferation of
CC  T cells or natural killer cells.
XX
SQ  Sequence 132 AA:

Query Match 35.3%; Score 6; DB 20; Length 132;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  4 SLHSS 9
    |||||
DB  48 slhss 53

RESULT 15
AAV30181
ID  AAV30181 standard; Protein: 132 AA.
XX
XX  AAV30181;
AC
XX
XX  29-OCT-1999 (first entry)
XX
XX  Amino acid sequence of the VL chain of rat antibody LO-CD2a.
XX
XX  Antibody LO-CD2a; CD2 antigen; T-lymphocyte; humanised antibody;
XX  T-cell-mediated immune response; graft rejection; autoimmune disease;
XX  graft-versus-host disease; T cell; natural killer cell.
XX
XX  Rattus sp.
XX
XX  US5951983-A.
XX
XX  14-SEP-1999.
XX
XX  07-JUN-1995; 95US-0477989.
XX

```

```

PR  07-JUN-1995; 95US-0477989.
PR  05-MAR-1993; 93US-0027008.
PR  09-SEP-1993; 93US-0119032.
PR  29-MAR-1995; 95US-0407009.
XX
XX  (BIOT-) BIO TRANSPLANT INC.
PA  (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX
PI  Bazin H, Kaplan R, Kieber-Emmons T, Latime D, Postema CE;
PI  White-Scharf ME;
DR  MPI: 1999-526991/44.
DR  N-PSDB: AA210196.
XX
XX  Antibody mediated inhibition of T cell immune response
XX
XX  Example 6: Fig 29; 104pp; English.
XX
XX  The present sequence represents the VL chain of rat monoclonal antibody
CC  LO-CD2a. LO-CD2a binds to an epitope of a CD2 antigen T-lymphocytes. The
CC  specification describes humanised LO-CD2a antibodies. The humanized
CC  LO-CD2a antibody comprises the human constant regions, a light chain
CC  framework region derived from a human antibody, a heavy chain framework
CC  region derived from a human antibody, heavy and light chain
CC  complementarity determining regions (CDRs) of the non-human monoclonal
CC  antibody produced by the cell line deposited as ATCC HB11423. The
CC  humanised antibodies are used in a method for treating a patient to
CC  inhibit a T-cell-mediated immune response. The method is useful for the
CC  treatment or prevention of graft rejection and graft-versus-host disease,
CC  as well as in the treatment of autoimmune diseases which are mediated
CC  by the activation and proliferation of T cells or natural killer cells.
XX
SQ  Sequence 132 AA:

Query Match 35.3%; Score 6; DB 20; Length 132;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  4 SLHSS 9
    |||||
DB  48 slhss 53

```

Search completed: September 15, 2001, 12:47:08  
Job time: 158 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:44:50 : Search time 35.36 Seconds  
(without alignments)  
9.899 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_150

Perfect score: 17  
Sequence: 1 KXISLHSSKEKLRER 17

Scoring table:

Gapop 60.0, Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size: 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Issued\_Patents\_AA:\*  
1: /cgn1\_7/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn1\_7/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn1\_7/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn1\_7/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn1\_7/ptodata/1/1aa/PC/TUS.COMB.pep:\*  
6: /cgn1\_7/ptodata/1/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	41.2	52	3	US-08-630-916A-76
2	6	35.3	112	1	US-08-477-877B-87
3	6	35.3	112	1	US-08-477-877B-88
4	6	35.3	112	2	US-08-472-281A-87
5	6	35.3	112	2	US-08-472-281A-88
6	6	35.3	112	2	US-08-477-989B-87
7	6	35.3	112	2	US-08-477-989B-88
8	6	35.3	112	4	US-09-240-274-30
9	6	35.3	132	1	US-08-477-877B-84
10	6	35.3	132	1	US-08-477-877B-81
11	6	35.3	132	2	US-08-472-281A-84
12	6	35.3	132	2	US-08-472-281A-91
13	6	35.3	132	2	US-08-477-989B-84
14	6	35.3	132	2	US-08-477-989B-91
15	6	35.3	292	2	US-08-928-284-2
16	6	35.3	866	1	US-08-100-692-1
17	6	35.3	866	1	US-08-674-030-1
18	6	35.3	874	2	US-08-247-904B-8
19	6	35.3	874	2	US-08-767-942A-21
20	5	29.4	6	5	PCR-US93-12679-5
21	5	29.4	16	1	US-08-438-123-1
22	5	29.4	18	1	US-08-438-123-9
23	5	29.4	21	1	US-08-704-170-28
24	5	29.4	21	1	US-08-383-753-70
25	5	29.4	21	2	US-08-586-772-70
26	5	29.4	21	2	US-08-959-512-70
27	5	29.4	21	5	PCR-US94-02631-28

28	5	29.4	26	2	US-08-620-151-114	Sequence 114, App
29	5	29.4	38	1	US-08-176-500-65	Sequence 65, Appl
30	5	29.4	38	1	US-08-471-052A-65	Sequence 65, Appl
31	5	29.4	38	1	US-08-189-331-65	Sequence 65, Appl
32	5	29.4	38	2	US-08-471-939-65	Sequence 65, Appl
33	5	29.4	38	2	US-08-471-800-65	Sequence 65, Appl
34	5	29.4	40	1	US-08-471-068-65	Sequence 65, Appl
35	5	29.4	40	2	US-08-190-802A-197	Sequence 197, App
36	5	29.4	53	3	US-08-630-916A-83	Sequence 83, Appl
37	5	29.4	78	2	US-08-343-443B-10	Sequence 10, Appl
38	5	29.4	79	2	US-08-611-510-6	Sequence 9, Appl
39	5	29.4	85	1	US-07-667-276A-9	Sequence 31, Appl
40	5	29.4	94	1	US-08-519-777-31	Sequence 31, Appl
41	5	29.4	94	1	US-08-742-035-31	Sequence 31, Appl
42	5	29.4	94	2	US-08-777-019-31	Sequence 31, Appl
43	5	29.4	94	2	US-08-777-143-31	Sequence 31, Appl
44	5	29.4	94	3	US-08-775-414-31	Sequence 31, Appl
45	5	29.4	94	4	US-08-931-858E-31	Sequence 31, Appl

#### ALIGNMENTS

RESULT 1  
US-08-630-916A-76  
; Sequence 76, Application US/08630916A  
; Patent No. 6011137  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio.  
; APPLICANT: Kay, Brian K.  
; APPLICANT: Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/630,916A  
; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18, 872  
; REFERENCE/DOCKET NUMBER: 1101-203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 896-8864/9741  
; INFORMATION FOR SEQ ID NO: 76:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 52 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-630-916A-76

Query Match 41.2%; Score 7; DB 3; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 SSKERLR 14  
|||||||

DB 30 SSKELR 36

RESULT 2  
US-08-477-877B-87  
Sequence 87, Application US/08477877B  
Patent No. 5730979  
GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Active  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,877B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995  
APPLICATION NUMBER: 08/119,032  
FILING DATE: 09-SEP-1993  
APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-146  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
FEATURE:  
NAME/KEY: Rat LO-CD2a light chain variable region.  
US-08-477-877B-87

Query Match: 35.3%; Score 6; DB 1; Length 112;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLHSS 9  
|||||  
DB 28 SLHSS 33

RESULT 3  
US-08-477-877B-88  
Sequence 88, Application US/08477877B  
Patent No. 5730979  
GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Active  
NUMBER OF SEQUENCES: 96

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
ADDRESSEE: Cecchi, Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,877B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995  
APPLICATION NUMBER: 08/119,032  
FILING DATE: 09-SEP-1993  
APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-146  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
FEATURE:  
NAME/KEY: Humanized LO-CD2a light chain variable region.  
US-08-477-877B-88

Query Match: 35.3%; Score 6; DB 1; Length 112;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLHSS 9  
|||||  
DB 28 SLHSS 33

RESULT 4  
US-08-472-281A-87  
Sequence 87, Application US/08472281A  
Patent No. 5817311  
GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Act  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
ADDRESSEE: Cecchi, Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS

SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,281A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995  
APPLICATION NUMBER: 08/119,032  
FILING DATE: 09-SEP-1993  
APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-142  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
FEATURE:  
NAME/KEY: Rat LO-CD2a light chain variable region.  
US-08-472-281A-87

Query Match 35.3%; Score 6; DB 2; Length 112;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLHSS 9  
|||||  
DB 28 SLHSS 33

RESULT 5  
US-08-472-281A-88  
Sequence 88, Application US/08472281A  
Patent No. 5817311  
GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gillfillan,  
ADDRESSSEE: Cecchi, Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,281A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995  
APPLICATION NUMBER: 08/119,032  
FILING DATE: 09-SEP-1993  
APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-142  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
FEATURE:  
NAME/KEY: Humanized LO-CD2a light chain variable region.  
US-08-472-281A-88

Query Match 35.3%; Score 6; DB 2; Length 112;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLHSS 9  
|||||  
DB 28 SLHSS 33

RESULT 6  
US-08-477-989B-87  
Sequence 87, Application US/08477989B  
Patent No. 5951983  
GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation and Proliferation  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gillfillan,  
ADDRESSSEE: Cecchi, Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,989B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995  
APPLICATION NUMBER: 08/119,032  
FILING DATE: 09-SEP-1993  
APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-147  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
FEATURE:  
NAME/KEY: Rat LO-CD2a light chain variable  
US-08-477-989B-87

Query Match 35.3%; Score 6; DB 2; Length 112;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLHSS 9  
|||||  
Db 28 SLHSS 33

RESULT 7  
US-08-477-989B-88  
Sequence 88, Application US/08477989B  
Patent No. 5951983  
GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
APPLICANT: Latine, Dominique  
APPLICANT: Kaplan, Ruth  
APPLICANT: Kiebler-Emmons, Thomas  
APPLICANT: Postema, Christina E.  
APPLICANT: White-Scharf, Mary  
TITLE OF INVENTION: LO-CD2a Antibody and Uses  
TITLE OF INVENTION: T-cell Activation and  
TITLE OF INVENTION: Proliferation  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08477,989B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995  
APPLICATION NUMBER: 08/119,032  
FILING DATE: 09-SEP-1993  
APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olsteid, Elliot M  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-147  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:

LENGTH: 112 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
FEATURE:  
NAME/KEY: Humanized LO-CD2a light chain  
US-08-477-989B-88

Query Match 35.3%; Score 6; DB 2; Length 112;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLHSS 9  
|||||  
Db 28 SLHSS 33

RESULT 8  
US-09-240-274-30  
Sequence 30, Application US/09240274  
Patent No. 6255455  
GENERAL INFORMATION:  
APPLICANT: Siegel, Donald L.  
TITLE OF INVENTION: RH(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL  
FILE REFERENCE: 09596-4202  
CURRENT APPLICATION NUMBER: US/09/240,274  
EARLIER FILING DATE: 1999-01-29  
EARLIER APPLICATION NUMBER: 60/081,380  
EARLIER FILING DATE: 1998-04-10  
EARLIER APPLICATION NUMBER: 60/028,550  
EARLIER FILING DATE: 1996-10-11  
NUMBER OF SEQ ID NOS: 224  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 30  
LENGTH: 112  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: anti-Rh(D) chain G01  
US-09-240-274-30

Query Match 35.3%; Score 6; DB 4; Length 112;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLHSS 9  
|||||  
Db 27 SLHSS 32

RESULT 9  
US-08-477-877B-84  
Sequence 84, Application US/08477877B  
Patent No. 5730979  
GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
APPLICANT: Latine, Dominique  
TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Act  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068  
COMPUTER READABLE FORM:



MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,877B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995  
APPLICATION NUMBER: 08/119,032  
FILING DATE: 09-SEP-1993  
APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-146  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 132 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
FEATURE:  
NAME/KEY: Chimeric LO-CD2a VL Chain  
US-08-477-877B-84

Query Match 35.3%, Score 6; DB 1; Length 132;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLHSS 9  
DB 48 SLHSS 53

RESULT 10  
US-08-477-877B-91  
Sequence 91, Application US/08477877B  
Patent No. 5730979  
GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Active  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,877B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995  
APPLICATION NUMBER: 08/119,032

FILING DATE: 09-SEP-1993  
APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-146  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 132 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
FEATURE:  
NAME/KEY: Humanized LO-CD2a light chain variable region.  
US-08-477-877B-91

Query Match 35.3%, Score 6; DB 1; Length 132;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLHSS 9  
DB 48 SLHSS 53

RESULT 11  
US-08-472-281A-84  
Sequence 84, Application US/08472281A  
Patent No. 5817311  
GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Act  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,281A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995  
APPLICATION NUMBER: 08/119,032  
FILING DATE: 09-SEP-1993  
APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-142  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:

LENGTH: 132 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
FEATURE:  
NAME/KEY: Chimeric LO-CD2a VL Chain  
US-08-472-281A-84

Query Match 35.3%; Score 6; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLHSS 9  
|||||  
DB 48 SLHSS 53

## RESULT 12

US-08-472-281A-91  
Sequence 91, Application US/08472281A  
Patent No. 5817311

GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
TITLE OF INVENTION: Latime, Dominique  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
STREET: Cecchi, Stewart & Olstein  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,281A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995

APPLICATION NUMBER: 08/119,032  
FILING DATE: 09-SEP-1993

APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.

REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-142

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 132 amino acids  
TYPE: amino acid

STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide

FEATURE:  
NAME/KEY: Humanized LO-CD2a light chain variable region.

US-08-472-281A-91

Query Match 35.3%; Score 6; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 31;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 SLHSS 9  
|||||  
DB 48 SLHSS 53

## RESULT 13

US-08-477-989B-84  
Sequence 84, Application US/08477989B  
Patent No. 5951983

GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
APPLICANT: Latime, Dominique  
APPLICANT: Kaplan, Ruth  
APPLICANT: Kieber-Emmons, Thomas  
APPLICANT: Postema, Christina E.  
APPLICANT: White-Scharf, Mary  
TITLE OF INVENTION: LO-CD2a Antibody and Uses  
TITLE OF INVENTION: Thereof for Inhibiting  
TITLE OF INVENTION: T-cell Activation and  
TITLE OF INVENTION: Proliferation  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
STREET: Cecchi, Stewart & Olstein  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,989B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995

APPLICATION NUMBER: 08/119,032  
FILING DATE: 09-SEP-1993

APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.

REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-147

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 132 amino acids  
TYPE: amino acid

STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide

FEATURE:  
NAME/KEY: Chimeric LO-CD2a VL Chain

US-08-477-989B-84

Query Match 35.3%; Score 6; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLHSS 9  
|||||  
DB 48 SLHSS 53

RESULT 14  
US-08-477-989B-91  
Sequence 91, Application US/08477989B  
Patent No. 5951983  
GENERAL INFORMATION:  
APPLICANT: Bazin, Herv  
APPLICANT: Latime, Dominique  
APPLICANT: Kaplan, Ruth  
APPLICANT: Kieber-Emmons, Thomas  
APPLICANT: Postema, Christina E.  
APPLICANT: White-Scharf, Mary  
TITLE OF INVENTION: LO-CD2a Antibody and Uses  
TITLE OF INVENTION: thereof for inhibiting  
TITLE OF INVENTION: T-Cell Activation and  
TITLE OF INVENTION: Proliferation  
NUMBER OF SEQUENCES: 96  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carrella, Byrne, Bain, Gilfillan,  
ADDRESSEE: Cecchi, Stewart & Olstein  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,989B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,009  
FILING DATE: 29-MAR-1995  
APPLICATION NUMBER: 08/119,032  
FILING DATE: 09-SEP-1993  
APPLICATION NUMBER: 08/027,008  
FILING DATE: 05-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Olstein, Elliot M.  
REGISTRATION NUMBER: 24,025  
REFERENCE/DOCKET NUMBER: 61750-147  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1744  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 132 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: polypeptide  
FEATURE:  
NAME/KEY: Humanized LO-CD2a light chain  
NAME/KEY: variable region.  
US-08-477-989B-91

Query Match 35.3%, Score 6; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLHSS 9  
|||||  
DB 48 SLHSS 53

RESULT 15  
US-08-928-284-2

Sequence 2, Application US/08928284  
Patent No. 5882871  
GENERAL INFORMATION:  
APPLICANT: Burnham, Martin K.  
TITLE OF INVENTION: No. 5882871el Saliva Binding Prot  
TITLE OF INVENTION: eln  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert, Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2793  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,284  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/026,682  
FILING DATE: 24-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Dickinson, O. Todd  
REGISTRATION NUMBER: 28,354  
REFERENCE/DOCKET NUMBER: P50543  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215/994-2252  
TELEFAX: 215/994-2222  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 292 amino acids  
TYPE: amino acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-928-284-2

Query Match 35.3%, Score 6; DB 2; Length 292;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRISLL 6  
|||||  
DB 2 KRISLL 7

Search completed: September 15, 2001, 12:47:51  
Job time: 181 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: September 15, 2001, 12:45:10 ; Search time 45.39 Seconds

(without alignments)  
28.530 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_150

Perfect score: 17

Sequence: 1 KRISLHSSKEKLRER 17

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database : PIR.68:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	7	41.2	250	2	H85067	hypothetical prote
2	7	41.2	486	2	F64204	spermidine/putresc
3	7	41.2	671	2	D38490	maturase-related p
4	7	41.2	673	2	S53604	maturase-related p
5	7	41.2	674	2	T07017	maturase homolog -
6	7	41.2	678	2	B38489	maturase-related p
7	7	41.2	687	2	A34879	maturase-related p
8	7	41.2	688	2	S53605	maturase-related p
9	7	41.2	716	2	G01627	androgen receptor
10	7	41.2	770	1	A44337	kinesin-related pr
11	7	41.2	791	2	T20815	hypothetical prote
12	7	41.2	1341	2	S66835	probable membrane
13	6	35.3	42	2	T07291	hypothetical prote
14	6	35.3	68	1	IABY3	proteolase A inhib
15	6	35.3	90	2	H69159	hypothetical prote
16	6	35.3	94	2	T12719	hypothetical prote
17	6	35.3	100	2	PH1075	Ig light chain V r
18	6	35.3	109	2	PH0089	Ig kappa chain V r
19	6	35.3	120	2	C29775	Ig kappa chain pre
20	6	35.3	135	2	T19002	hypothetical prote
21	6	35.3	159	2	G84974	transcription elon
22	6	35.3	160	2	T24339	hypothetical prote
23	6	35.3	160	2	S52889	probable membrane
24	6	35.3	163	2	D64025	hypothetical prote
25	6	35.3	169	2	I40131	exported protein A
26	6	35.3	177	2	E71138	probable phosphori
27	6	35.3	181	1	JVBPNL	DNA-packaging prot
28	6	35.3	181	2	F64788	DNA packaging prot
29	6	35.3	181	2	G85689	hypothetical prote

30	6	35.3	184	2	T38315	very hypothetical
31	6	35.3	189	2	G64909	DNA packaging prot
32	6	35.3	207	2	T25601	hypothetical prote
33	6	35.3	218	2	C75394	ATP phosphoribosyl
34	6	35.3	219	2	F71155	hypothetical prote
35	6	35.3	222	2	S71231	geranylgeranyl pyr
36	6	35.3	236	2	A40143	placental lactogen
37	6	35.3	236	2	A37930	hypothetical prote
38	6	35.3	239	2	A84091	hypothetical prote
39	6	35.3	242	2	A75023	hypothetical prote
40	6	35.3	242	2	S39642	motility protein h
41	6	35.3	249	2	B75110	lipote-protein 1i
42	6	35.3	263	2	B70153	conserved hypothet
43	6	35.3	265	2	S34668	nodulation protein
44	6	35.3	281	2	T24658	hypothetical prote
45	6	35.3	282	2	D81326	5,10-methyleneetr

## ALIGNMENTS

RESULT 1  
H85067 hypothetical protein ATG05400 (imported) - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001  
C:Accession: H85067  
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488  
A:Accession: H85067  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-250 <STO>  
A:Cross-references: GB:NC\_001268; NID:g7267300; PID:CA881082.1; GSPDB:GN00140  
A:Gene: ATG05400  
A:Map position: 4  
C:Superfamily: Arabidopsis thaliana hypothetical protein F7J7.80

Query Match 41.2%; Score 7; DB 2; Length 250;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 SSKELR 14  
DB 244 SSKELR 250

RESULT 2  
F64204 spermidine/putrescine transport ATP-binding protein potA homolog - Mycoplasma genital  
C:Species: Mycoplasma genitalium  
C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 17-Mar-2000  
C:Accession: F64204  
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.  
M.; Fuhmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,  
C.A.; Venter, J.C.  
Science 270, 397-403, 1995  
A:Title: The minimal gene complement of Mycoplasma genitalium.  
A:Reference number: A64200; MUID:96026346  
A:Accession: F64204  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-486 <TIGR>  
A:Cross-references: GB:U39683; GB:I43967; NID:g1045711; PID:g1045714; TIGR:MG042  
A:Experimental source: Strain G-37  
C:Genetics:  
A:Genetic code: SGC3  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog  
C:Keywords: ATP

F:1-351/Domain: ATP-binding cassette, homology #status atypical <ABCL>

Query Match 41.2% Score 7; DB 2; Length 486;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 KEKLRRE 16  
|||||  
DB 100 KEKLRRE 106

RESULT 3

D38490  
maturase-related protein (nad1 intron) - evening primrose mitochondrion  
C:Species: mitochondrion Oenothera villaricae (evening primrose)  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 30-Sep-1993  
C:Accession: D38490; B38490  
R:Wissinger, B.; Schuster, W.; Brennicke, A.  
Cell 65, 473-482, 1991  
A:Title: Trans splicing in Oenothera mitochondria: nad1 mRNAs are edited in exon and tra  
A:Reference number: A38490; MUID:91208684  
A:Accession: D38490  
A:Molecule type: mRNA  
A:Residues: 1-671 <MIS>  
A:Cross-references: GB:M63034  
A:Accession: B38490  
A:Molecule type: DNA  
A:Residues: 1,'S',3-5,'P',7-26,'S',28-30,'P',32-80,'P',82-100,'S',102-575,'S',577-582,'F  
A:Cross-references: GB:M63034  
C:Genetics:  
A:Genome: mitochondrion  
C:Keywords: mitochondrion; RNA editing

Query Match 41.2% Score 7; DB 2; Length 671;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9  
|||||  
DB 475 ISLHSS 481

RESULT 4

S53604  
maturase-related protein (nad1 intron) - maize mitochondrion (fragment)  
C:Species: mitochondrion Zea mays (maize)  
C>Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 29-Oct-1999  
C:Accession: S53604  
R:Thomson, M.C.; Macfarlane, J.L.; Beagley, C.T.; Wolstenholme, D.R.  
Nucleic Acids Res. 22, 5745-5752, 1994  
A:Title: RNA editing of mat-r transcripts in maize and soybean increases similarity of t  
protein.  
A:Reference number: S53604; MUID:95140641  
A:Accession: S53604  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-673 <THO>  
A:Cross-references: GB:U09987; NID:g607797; PIDN:AA67713.1; PID:g840925  
C:Genetics:  
A:Genome: mitochondrion  
C:Keywords: mitochondrion; RNA editing

Query Match 41.2% Score 7; DB 2; Length 673;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9  
|||||  
DB 476 ISLHSS 482

RESULT 5

T07017  
maturase homolog - potato mitochondrion (fragment)  
C:Species: mitochondrion Solanum tuberosum (potato)  
C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 21-Jul-2000  
C:Accession: T07017  
R:Bequ, D.; Mercado, A.; Farre, J.C.; Moenne, A.; Holuigue, L.; Araya, A.; Jordana, X  
Curr. Genet. 33, 420-428, 1998  
A:Title: Editing status of mat-r transcripts in mitochondria from two plant species:  
A:Reference number: Z15865; MUID:96309836

A:Accession: T07017  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-674 <BEQ>  
A:Cross-references: EMBL:AJ003130; NID:g3336907; PIDN:CA05884.1; PID:g3336908  
A:Experimental source: cv. Bintje; tuber  
C:Genetics:  
A:Gene: mat-r  
A:Genome: mitochondrion  
C:Keywords: mitochondrion; RNA editing

Query Match 41.2% Score 7; DB 2; Length 674;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9  
|||||  
DB 475 ISLHSS 481

RESULT 6

B38489  
maturase-related protein (nad1 intron) - wheat mitochondrion  
C:Species: mitochondrion Triticum aestivum (common wheat)  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 29-Oct-1999  
C:Accession: B38489  
R:Chappelaine, Y.; Bonen, L.  
Cell 65, 465-472, 1991  
A:Title: The wheat mitochondrial gene for subunit I of the NADH dehydrogenase complex  
A:Reference number: A38489; MUID:91208683  
A:Accession: B38489  
A:Molecule type: DNA  
A:Residues: 1-678 <CHA>  
A:Cross-references: EMBL:X57965; NID:g14252; PIDN:CAA4103.1; PID:el57959; PID:g13346  
A>Note: this reading frame extends between two stop codons and does not begin with a  
C:Genetics:  
A:Genome: mitochondrion  
C:Keywords: mitochondrion

Query Match 41.2% Score 7; DB 2; Length 678;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9  
|||||  
DB 481 ISLHSS 487

RESULT 7

A34879  
maturase-related protein (nad1 intron) - fava bean mitochondrion  
C:Species: mitochondrion Vicia faba (fava bean)  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 30-Sep-1993  
C:Accession: A34879  
R:Walleitner, J.A.; Macfarlane, J.L.; Wolstenholme, D.R.  
Proc. Natl. Acad. Sci. U.S.A. 87, 548-552, 1990  
A:Title: A sequence encoding a maturase-related protein in a group II intron of a pla  
A:Reference number: A34879; MUID:90138867  
A:Accession: A34879  
A:Molecule type: DNA

A:Residues: 1-687 <MAH>  
 A:Cross-references: EMBL:M30176  
 A:Note: the authors translated the codon CCG for residues 51, 210, 367, 378, 398, 485, 5  
 C:Genetics:  
 A:Genome: mitochondrion  
 C:Keywords: mitochondrion

Query Match 41.2%; Score 7; DB 2; Length 687;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9  
 Db 477 ISLHSS 483

RESULT 8  
 S53605  
 maturase-related protein (nad1 intron) - soybean mitochondrion (fragment)  
 C:Species: mitochondrion Glycine max (soybean)  
 C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 21-Jul-2000  
 C:Accession: S53605  
 R:Thomson, M.C.; Macfarlane, J.L.; Beagley, C.T.; Wolstenholme, D.R.  
 Nucleic Acids Res. 22, 5745-5752, 1994  
 A:Title: RNA editing of mat-r transcripts in maize and soybean increases similarity of t  
 protein.  
 A:Reference number: S53604; MUID:95140641  
 A:Accession: S53605  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-688 <THO>  
 A:Cross-references: GB:U09988; NID:9607800; PIDN:AAA67711.1; PID:9840924  
 C:Genetics:  
 A:Genome: mitochondrion  
 C:Keywords: mitochondrion; RNA editing

Query Match 41.2%; Score 7; DB 2; Length 688;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 ISLHSS 9  
 Db 478 ISLHSS 484

RESULT 9  
 G01627  
 androgen receptor 1 - human  
 N:Alternate names: dihydrotestosterone receptor  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 17-Jul-1998  
 C:Accession: G01627  
 R:Blum, M.  
 submitted to the EMBL Data Library, December 1994  
 A:Reference number: G07955  
 A:Accession: G01627  
 A:Status: Preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: mRNA  
 A:Residues: 1-716 <BIN>  
 A:Cross-references: EMBL:U19345; NID:g726041; PID:g726042  
 C:Genetics:  
 A:Gene: GDB:AR; DHTR; SBMA  
 A:Cross-references: GDB:120556; OMIM:313700  
 A:Map position: Xq11-Xq12

Query Match 41.2%; Score 7; DB 2; Length 716;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LTHSKE 11

Db 120 LTHSKE 126

RESULT 10  
 A44337  
 kinesin-related protein KLPa - Emericella nidulans  
 N:Alternate names: kinesin-like protein, KAR3-related; KLPa protein  
 C:Species: Emericella nidulans, Aspergillus nidulans  
 C:Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 19-Jan-2001  
 C:Accession: A44337; S24830  
 R:O'Connell, M.J.; Meluh, P.B.; Rose, M.D.; Morris, N.R.  
 J. Cell Biol. 120, 153-162, 1993  
 A:Title: Suppression of the bimC mitotic spindle defect by deletion of klpA, a gene  
 A:Reference number: A44337; MUID:93107178  
 A:Accession: A44337  
 A:Molecule type: mRNA  
 A:Residues: 1-770 <OLC>  
 A:Cross-references: GB:X64603; NID:g2703; PIDN:CAA5887.1; PID:g2704  
 A:Note: sequence extracted from NCBI backbone (NCBIP:121121)  
 C:Genetics:  
 A:Map position: 1  
 A:Gene: KLPa  
 C:Superfamily: Kinesin-related protein KLPa; kinesin motor domain homology  
 C:Keywords: Arp; coiled coil; microtubule binding; nucleotide binding; P-loop  
 F:422-762/Domain: kinesin motor domain homology <KMT>  
 F:514-521/Region: nucleotide-binding motif A (P-loop)  
 F:520/Binding site: Arp (Lys) #status predicted

Query Match 41.2%; Score 7; DB 1; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 KEKLRE 16  
 Db 398 KEKLRE 404

RESULT 11  
 T20815  
 hypothetical protein F26D2.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T20815; T21409  
 R:McMurray, A.  
 submitted to the EMBL Data Library, March 1997  
 A:Reference number: Z19327  
 A:Accession: T20815  
 A:Status: Preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-791 <WIL>  
 A:Cross-references: EMBL:Z93377; PIDN:CAB07581.1; GSPDB:GN00023; CESP:F26D2.2  
 A:Experimental source: clone F13A7  
 R:McMurray, A.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z19418  
 A:Accession: T21409  
 A:Status: Preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-791 <WIL>  
 A:Cross-references: EMBL:Z81513; PIDN:CAB04186.1; GSPDB:GN00023; CESP:F26D2.2  
 A:Experimental source: clone F26D2  
 C:Genetics:  
 A:Gene: CESP:F26D2.2  
 A:Map position: 5  
 A:Introns: 28/3; 194/2; 195/3; 299/2; 404/1; 462/3; 500/1; 670/3; 698/3

Query Match 41.2%; Score 7; DB 2; Length 791;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 EXLRER 17  
|||||  
Db 153 EXLRER 159

## RESULT 12

S66835  
probable membrane protein YOL138c - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: hypothetical protein AOD1341; hypothetical protein 00483  
C:Species: *Saccharomyces cerevisiae*  
C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 21-Jul-2000  
C:Accession: S66835; S72030  
R:Arino, J.; Casamayor, A.; Gamo, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas, C.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S66814  
A:Accession: S66835  
A:Molecule type: DNA  
A:Residues: 1-1341 <ARI>  
A:Cross-references: EMBL:Z74880; NID:q1420031; PID:e251919; PID:q1420032; MIPS:YOL138c  
A:Experimental source: strain S288C  
R:Aldea, M.; Piedraflita, L.; Casas, C.; Casamayor, A.; Khalid, H.; Balcells, L.; Arino, J.  
Yeast 12, 1053-1058, 1996  
A:Title: Sequence analysis of a 12 801 bp fragment of the left arm of yeast chromosome X  
protein and six other open reading frames.  
A:Reference number: S72030; MUID:97051593  
A:Accession: S72030  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1223, 'V', 1224-1341 <ALD>  
A:Cross-references: EMBL:X95465; NID:q1628437; PIDN:CAA64732.1; PID:q1628438  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996  
C:Genetics:  
A:Map position: 15L  
A:Note: YOL138c  
C:Superfamily: unassigned WD repeat proteins; WD repeat homology  
C:Keywords: transmembrane protein  
F:254-288/Domain: WD repeat homology <WD1>  
F:365-398/Domain: WD repeat homology <WD2>  
F:1178-1194/Domain: transmembrane #status predicted <TMW>

Query Match 41.2%; Score 7; DB 2; Length 1341;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SSKELR 14  
|||||  
Db 1257 SSKELR 1263

## RESULT 13

T07291  
hypothetical protein 42c - *Chlorella vulgaris* chloroplast  
C:Species: *Chlorella vulgaris*  
C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000  
C:Accession: T07291  
R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakae  
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga *Chl*  
A:Reference number: Z15985; MUID:97303241  
A:Accession: T07291  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-42 <MAK>  
A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BA57939.1; PID:g2224455  
C:Genetics:  
A:Genome: chloroplast  
A:Keywords: chloroplast

Query Match 35.3%; Score 6; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKISLL 6  
|||||  
Db 34 KKISLL 39

## RESULT 14

IABY3  
protease A inhibitor 3 - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: protein YMB010.04c; protein YMR174c  
C:Species: *Saccharomyces cerevisiae*  
C:Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 21-Jul-2000  
C:Accession: A01334; S16692; S55121  
R:Biedermaier, K.; Montali, U.; Martin, B.; Svendsen, I.; Ottesen, M.  
Carlsberg Res. Commun. 45, 225-235, 1980  
A:Title: The amino acid sequence of proteinase A inhibitor 3 from baker's yeast.  
A:Reference number: A01334  
A:Accession: A01334  
A:Molecule type: protein  
A:Residues: 1-68 <BIE>  
A:Note: nearly all of the inhibitory activity is present in the peptide consisting of  
R:Schu, P.; Wolf, D.H.  
FEBS Lett. 283, 78-84, 1991  
A:Title: The proteinase yscA-inhibitor 1(A)(3), gene. Studies of cytoplasmic protein  
A:Reference number: S16692; MUID:91243884  
A:Accession: S16692  
A:Molecule type: DNA  
A:Residues: 1-68 <SCH>  
A:Cross-references: EMBL:X60050; NID:q4094; PIDN:CAA42650.1; PID:q4095  
R:Churcher, C.M.  
submitted to the EMBL Data Library, June 1995  
A:Reference number: S55118  
A:Accession: S55121  
A:Molecule type: DNA  
A:Residues: 1-68 <CHD>  
A:Cross-references: EMBL:Z49808; NID:9854440; PIDN:CAA89907.1; PID:9854444; GSPDB:GNO  
A:Experimental source: strain AB972  
C:Genetics:  
A:Gene: SGD:PAI3; MIPS:YMR174c  
A:Cross-references: SGD:S0004786; MIPS:YMR174c  
A:Map position: 13R  
C:Superfamily: proteinase A inhibitor 3  
C:Keywords: acetylated amino end; blocked amino end; proteinase inhibitor  
F:1/Modified site: acetylated amino end (Met) #status experimental

Query Match 35.3%; Score 6; DB 1; Length 68;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SSKEL 13  
|||||  
Db 14 SSKEL 19

## RESULT 15

H69159  
hypothetical protein MTH458 - *Methanobacterium thermoautotrophicum* (strain Delta H)  
C:Species: *Methanobacterium thermoautotrophicum*  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: H69159  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.  
Qiu, D.; Spadafora, R.; Vicalire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwanti,  
K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7151, 1997  
A:Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: fu  
A:Reference number: A69000; MUID:98037514  
A:Accession: H69159  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-90 <MTH>  
A:Cross-references: GB:AE000830; GB:AE000666; NID:g2621523; PIDN:AMB84964.1; PID:g262  
A:Experimental source: strain Delta H



C;Genetics:  
A;Gene: MTH458

Query Match	35.38;	Score 6;	DB 2;	Length 90;
Best Local Similarity	100.08;	Pred. No. 25;		
Matches	6;	Conservative 0;	Mismatches 0;	Indels 0;
				Gaps 0;

OY	10	KEKLR	15
Db	36	KEKLR	41

```
Search completed: September 15, 2001, 12:48:49
Job time: 219 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:50:26 : Search time 17.5 Seconds  
(without alignments)  
128.525 Million cell updates/sec

Title: us-09-389-000-2\_COPY\_134\_150  
Perfect score: 17  
Sequence: 1 KKISLHSSKEKLRER 17

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size: 0

Total number of hits satisfying chosen parameters: 105709

Minimum DB seq length: 0

Maximum DB seq length: 116

Post-processing: Listing first 45 summaries

Database:

SPTREMBL\_16:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mhc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rodent:  
12: sp\_unclassified:  
13: sp\_vertebrate:  
14: sp\_virus:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	35.3	39	4	Q9UBN9
2	6	35.3	42	8	O20166
3	6	35.3	90	1	O26558
4	6	35.3	94	1	O9HB63
5	6	35.3	94	9	O80193
6	5	29.4	39	2	P74773
7	5	29.4	42	8	O20145
8	5	29.4	43	2	O9K157
9	5	29.4	46	10	O9ESV1
10	5	29.4	47	8	O20181
11	5	29.4	47	14	O9OCM2
12	5	29.4	47	14	O9OCL9
13	5	29.4	47	14	O9OCL6
14	5	29.4	47	14	O9OCL3
15	5	29.4	47	14	O9OCL0
16	5	29.4	47	14	O9OCK7
17	5	29.4	47	14	O9OCK4
18	5	29.4	47	14	O9OCK1
19	5	29.4	53	2	O9PGC6

20	5	29.4	53	14	P88748	P88748 human immun
21	5	29.4	55	4	O9UGF2	O9UGF2 homo sapien
22	5	29.4	59	2	O44673	O44673 borrelia af
23	5	29.4	59	2	O57397	O57397 borrelia af
24	5	29.4	59	10	O38962	O38962 arabidopsis
25	5	29.4	62	2	O24728	O24728 shigella fl
26	5	29.4	62	4	O95690	O95690 homo sapien
27	5	29.4	62	4	O9HB84	O9HB84 homo sapien
28	5	29.4	63	4	O9Y317	O9Y317 homo sapien
29	5	29.4	64	2	O50699	O50699 borrelia bu
30	5	29.4	64	2	O9EY25	O9EY25 photobacter
31	5	29.4	65	14	O97037	O97037 human immun
32	5	29.4	68	4	O9UHT6	O9UHT6 homo sapien
33	5	29.4	69	13	O91894	O91894 orthorhynch
34	5	29.4	71	1	O58638	O58638 methanococ
35	5	29.4	72	14	O9DSV5	O9DSV5 ascovirus d
36	5	29.4	74	14	O91HR5	O91HR5 tl virus. o
37	5	29.4	78	2	O92LP5	O92LP5 helicobacte
38	5	29.4	79	2	O47383	O47383 escherichia
39	5	29.4	79	4	O95689	O95689 homo sapien
40	5	29.4	79	4	O9U176	O9U176 homo sapien
41	5	29.4	81	2	O9EZM9	O9EZM9 wolbachia e
42	5	29.4	81	14	O9WNE1	O9WNE1 human immun
43	5	29.4	82	14	O9WNE2	O9WNE2 human immun
44	5	29.4	83	4	O9NU02	O9NU02 homo sapien
45	5	29.4	84	2	O49311	O49311 mycoplasma

## ALIGNMENTS

RESULT 1  
Q9UBN9 PRELIMINARY: PRT: 39 AA.  
AC Q9UBN9;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE E6-AP UBIQUITIN-PROTEIN LIGASE (FRAGMENT).  
GN UBE3A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE=98126441; PubMed=9465301;  
RA Kishino T., Wagstaff J.;  
RT "Genomic organization of the UBE3A/E6-AP gene and related  
RT pseudogenes.";  
RL Genomics 47:101-107(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Hennies H.C., Buerger J., Sperling K., Reis A.;  
RT "Mutations in the E6-AP gene (UBE3A) in patients with Angelman  
RT syndrome.";  
RL Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF009341; AAC39580.1; -;  
DR EMBL: AJ001113; CA04540.1; -;  
DR InterPro: IPR000569; -;  
DR PROSITE: PS50237; HECT; 1.  
KW Ligase.  
FT NON\_TER  
SQ SEQUENCE 39 AA; 4441 MW; 763722F37AFA7193 CRC64;

Query Match 35.3%, Score 6; DB 4; Length 39;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 8 SSKKLL 13  
|||||

Db 17 SKEKL 22

## RESULT 2

020166 PRELIMINARY; PRT: 42 AA.  
AC 020166;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE ORF42C.  
OS Chlorococcoides vulgaris.  
OC Chlorococcoides.  
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorococcoidales;  
OC Chlorococcoidaceae; Chlorococcoides.  
OX NCBI\_TaxID=3077;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97303241; PubMed=9159184;  
RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,  
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,  
RA Inamura A., Yoshinaga K., Sugita M.;  
RT "Complete nucleotide sequence of the chloroplast genome from the green  
RT alga Chlorococcoides vulgaris: the existence of genes possibly involved in  
RT chloroplast division."  
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
DR EMBL: AB001684; BAA57939.1; -.  
KW Chloroplast.  
SQ SEQUENCE 42 AA: 5355 MW: 136FF356DC528715 CRC64;

Query Match 35.3%; Score 6; DB 8; Length 42;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKSIL 6  
DB 34 KKSIL 39

RESULT 3  
026558 PRELIMINARY; PRT: 90 AA.  
AC 026558;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL 10.5 KDA PROTEIN.  
GN MTH438.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
OC Methanobacterium.  
OX NCBI\_TaxID=145262;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DELTA H;  
RX MEDLINE=98037514; PubMed=9371463;  
RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,  
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
RA Spadatoro R., Vitcare R., Wang Y., Wierzbowski J., Gibson R.,  
RA Jiwani N., Carno A., Bush D., Safer H., Patwell D., Prabhakar S.,  
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
RA Daniels C.J., Mao J.-I., Rice P., Nollan J., Reeve J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
RT delta: functional analysis and comparative genomics."  
RL J. Bacteriol. 179:7135-7155(1997).  
DR EMBL: AE000830; AAB84964.1; -.  
DR InterPro: IPR000014; -.  
DR InterPro: IPR000700; -.  
DR Pfam: PF00989; PAS; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 90 AA: 10547 MW: 1E3CAF1C206B9D7F CRC64;

Query Match 35.3%; Score 6; DB 1; Length 90;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 KEKLR 15  
DB 36 KEKLR 41

## RESULT 4

09HH63 PRELIMINARY; PRT: 94 AA.  
AC 09HH63;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL 11.5 KDA PROTEIN.  
OS Methanobacterium wolfei.  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
OC Methanobacterium.  
OX NCBI\_TaxID=145261;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Luo Y., Pfister P., Leisinger T., Wasserfallen A.;  
RT "The genome of archaeal prophage pSM100 encodes the lytic enzyme  
RT responsible for autolysis of Methanobacterium wolfei."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF01375; AAC39948.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 94 AA: 11511 MW: E3FFA7957137B7FC CRC64;

Query Match 35.3%; Score 6; DB 1; Length 94;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 SKEKL 14  
DB 72 SKEKL 77

RESULT 5  
080193 PRELIMINARY; PRT: 94 AA.  
AC 080193;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE HYPOTHETICAL 11.5 KDA PROTEIN.  
OS Methanobacterium phage pSM2.  
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.  
OX NCBI\_TaxID=77048;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Pfister P., Wasserfallen A., Stettler R., Leisinger T.;  
RT "Archaeophage pSM2 complete genomic DNA."  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF055411; AAC27042.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 94 AA: 11483 MW: E3FFA79572C478C3 CRC64;

Query Match 35.3%; Score 6; DB 9; Length 94;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 SKEKL 14  
DB 72 SKEKL 77

```

RESULT 6
P74773 PRELIMINARY; PRT; 39 AA.
AC P74773;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 4.7 KDA PROTEIN.
GN SGL0001.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RA Tabata S.;
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome."
RL DNA Res. 2:153-166(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naro K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions."
RL DNA Res. 3:109-136(1996).
DR EMBL; D64003; BAA10494.1; -.
KW Hypothetical protein.
SQ SEQUENCE 39 AA; 4659 MW; EBD9HE03539EAF2B CRC64;

Query Match 29.4%; Score 5; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KISL 6
Db 9 KISL 13

RESULT 7
O20145 PRELIMINARY; PRT; 42 AA.
AC O20145;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE ORF42A.
OS Chlorella vulgaris.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3077;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97303241; PubMed=9159184;
RA Takasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA Imanura A., Yoshinaga K., Sugitara M.;
RT "Complete nucleotide sequence of the chloroplast genome from the green
RT alga Chlorella vulgaris; the existence of genes possibly involved in

```

```

RT Chloroplast division."
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
DR EMBL; AB001684; BAA57895.1; -.
KW Chloroplast.
SQ SEQUENCE 42 AA; 5252 MW; 445B221534CACD71 CRC64;

Query Match 29.4%; Score 5; DB 8; Length 42;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 SREKL 13
Db 19 SREKL 23

RESULT 8
Q9K1S7 PRELIMINARY; PRT; 43 AA.
AC Q9K1S7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HYPOTHETICAL PROTEIN CP1078.
GN CP1078.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RA Read T.D., Brumham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Umayam L.A., Utterback T.,
RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
RA Bowman C., Dodson R., Gwin M., Nelson W., Deboy R., Kolonay J.,
RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AE002264; AAF38850.1; -.
KW Hypothetical protein.
SQ SEQUENCE 43 AA; 5140 MW; 7F9BAAPD6A313336 CRC64;

Query Match 29.4%; Score 5; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KISL 5
Db 7 KISL 11

RESULT 9
Q9FSV1 PRELIMINARY; PRT; 46 AA.
AC Q9FSV1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PROTEIN KINASE (FRAGMENT).
GN PKF4.
OS Fagus sylvatica (Beechnut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fagales; Fagaceae; Fagus.
OX NCBI_TaxID=28930;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RA Lorenzo O.;

```

RL Thesis (2000), Fisiologia Vegetal, Facultad de Biologia,  
 RL Universidad de Salamanca.  
 DR EMBL: AJ298982; CAC09570.1; -  
 KW Kinase  
 FT NON\_TER 1 1  
 FT NON\_TER 46 46  
 SQ SEQUENCE 46 AA; 5051 MW; C64FEF5748A98822 CRC64;

Query Match 29.4%; Score 5; DB 10; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLHS 8  
 DB 28 SLHS 32

RESULT 10  
 O20181  
 ID 020181 PRELIMINARY; PRT; 47 AA.  
 AC 020181.  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE ORF47B.  
 OS Chlorella vulgaris.  
 OC Chloroplast.  
 CC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
 OC Chlorellaceae; Chlorellia.  
 OX NCBI\_TaxID=3077;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97303241; PubMed=9159184;  
 RA Wakesugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,  
 RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,  
 RA Imamura A., Yoshinaga K., Sugitara M.,  
 RT "Complete nucleotide sequence of the chloroplast genome from the green  
 alga Chlorella vulgaris: the existence of genes possibly involved in  
 RT chloroplast division."  
 RT Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
 RL EMBL: AB001684; BAA57959.1; -  
 DR EMBL: AB001684; BAA57959.1; -  
 KM CHLOROPLAST.  
 SQ SEQUENCE 47 AA; 5867 MW; B161224E524F3802 CRC64;

Query Match 29.4%; Score 5; DB 8; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SSKER 12  
 DB 3 SSKER 7

RESULT 11  
 O90CM2  
 ID 090CM2 PRELIMINARY; PRT; 47 AA.  
 AC 090CM2.  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE P24 ORF (FRAGMENT).  
 OS Borna disease virus (BDV).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales.  
 OX NCBI\_TaxID=12455;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99445641; PubMed=10515835;  
 RA Czygan M., Hallensleben W., Hofer M., Pollak S., Sauder C., Bilzer T.,  
 RA Blumcke I., Riederer P., Bogerts B., Falkai P., Schwarz M.J.,  
 RA Maslian E., Staeheli P., Hufert F.T., Lieb K.;  
 RT "Borna disease virus in human brains with a rare form of hippocampal

RT degeneration but not in brains of patients with common  
 RT neuropsychiatric disorders."  
 RT J. Infect. Dis. 180:1695-1699(1999).  
 DR EMBL: AJ246858; CAB52089.1; -  
 KW Kinase  
 FT NON\_TER 1 1  
 FT NON\_TER 47 47  
 SQ SEQUENCE 47 AA; 5258 MW; 440FCD9B41B23F38 CRC64;

Query Match 29.4%; Score 5; DB 14; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LRRER 17  
 DB 21 LRRER 25

RESULT 12  
 O90CL9  
 ID 090CL9 PRELIMINARY; PRT; 47 AA.  
 AC 090CL9.  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE P24 ORF (FRAGMENT).  
 OS Borna disease virus (BDV).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales.  
 OX NCBI\_TaxID=12455;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99445641; PubMed=10515835;  
 RA Czygan M., Hallensleben W., Hofer M., Pollak S., Sauder C., Bilzer T.,  
 RA Blumcke I., Riederer P., Bogerts B., Falkai P., Schwarz M.J.,  
 RA Maslian E., Staeheli P., Hufert F.T., Lieb K.;  
 RT "Borna disease virus in human brains with a rare form of hippocampal  
 RT degeneration but not in brains of patients with common  
 RT neuropsychiatric disorders."  
 RT J. Infect. Dis. 180:1695-1699(1999).  
 DR EMBL: AJ246859; CAB52092.1; -  
 FT NON\_TER 1 1  
 FT NON\_TER 47 47  
 SQ SEQUENCE 47 AA; 5258 MW; 440FCD9B41B23F38 CRC64;

Query Match 29.4%; Score 5; DB 14; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LRRER 17  
 DB 21 LRRER 25

RESULT 13  
 O90CL6  
 ID 090CL6 PRELIMINARY; PRT; 47 AA.  
 AC 090CL6.  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE P24 ORF (FRAGMENT).  
 OS Borna disease virus (BDV).  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales.  
 OX NCBI\_TaxID=12455;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99445641; PubMed=10515835;  
 RA Czygan M., Hallensleben W., Hofer M., Pollak S., Sauder C., Bilzer T.,  
 RA Blumcke I., Riederer P., Bogerts B., Falkai P., Schwarz M.J.,  
 RA Maslian E., Staeheli P., Hufert F.T., Lieb K.;  
 RT "Borna disease virus in human brains with a rare form of hippocampal  
 RT degeneration but not in brains of patients with common

RT neuropsychiatric disorders.";  
RL J. Infect. Dis. 180:1695-1699(1999).  
DR EMBL: AJ246860; CAB52095.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 47 47  
SQ SEQUENCE 47 AA: 5258 MW: 440PCD9B41B23F38 CRC64;

Query Match 29.4%: Score 5; DB 14; Length 47;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LRRR 17  
|||||  
DB 21 LRRR 25

RESULT 14  
O9OCL3 PRELIMINARY; PRT; 47 AA.  
AC O9OCL3:  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE P24 ORF (FRAGMENT).  
OS Borna disease virus (BDV).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales.  
OX NCBI\_TaxID=12455;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99445641; PubMed=10515835;  
RA Czygan M., Hallensleben W., Hofer M., Pollak S., Sauder C., Bilzer T.,  
RA Blumcke I., Riederer P., Bogerts B., Falkai P., Schwarz M.J.,  
RA Masliah E., Staeheli P., Hufert F.T., Lieb K.;  
RT "Borna disease virus in human brains with a rare form of hippocampal  
RT degeneration but not in brains of patients with common  
RT neuropsychiatric disorders.";  
RL J. Infect. Dis. 180:1695-1699(1999).  
DR EMBL: AJ246861; CAB52098.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 47 47  
SQ SEQUENCE 47 AA: 5258 MW: 440PCD9B41B23F38 CRC64;

Query Match 29.4%: Score 5; DB 14; Length 47;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LRRR 17  
|||||  
DB 21 LRRR 25

RESULT 15  
O9OCL0 PRELIMINARY; PRT; 47 AA.  
AC O9OCL0:  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE P24 ORF (FRAGMENT).  
OS Borna disease virus (BDV).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales.  
OX NCBI\_TaxID=12455;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99445641; PubMed=10515835;  
RA Czygan M., Hallensleben W., Hofer M., Pollak S., Sauder C., Bilzer T.,  
RA Blumcke I., Riederer P., Bogerts B., Falkai P., Schwarz M.J.,  
RA Masliah E., Staeheli P., Hufert F.T., Lieb K.;  
RT "Borna disease virus in human brains with a rare form of hippocampal  
RT degeneration but not in brains of patients with common  
RT neuropsychiatric disorders.";

RL J. Infect. Dis. 180:1695-1699(1999).  
DR EMBL: AJ246862; CAB52101.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 47 47  
SQ SEQUENCE 47 AA: 5258 MW: 440PCD9B41B23F38 CRC64;

Query Match 29.4%: Score 5; DB 14; Length 47;  
Best Local Similarity 100.0%; Pred. No. 4.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LRRR 17  
|||||  
DB 21 LRRR 25

Search completed: September 15, 2001, 12:52:19  
Job time: 113 sec

---



Result No.	Query %	Score	Match	Length	DB	ID	Description
1	6	35.3	68	1	IPA3_YEAST	P01094	saccharomyc
2	5	29.4	61	1	DNE1_BFDV	P33893	budgerigar
3	5	29.4	62	1	YH45_CAEEL	P34521	caenorhabdi
4	5	29.4	63	1	YORQ_TTV1	P19301	thermoprote
5	5	29.4	68	1	RL29_PYRHO	O74094	pyrococcus
6	5	29.4	72	1	HST1_ECOLI	P01559	escherichia
7	5	29.4	72	1	HST2_ECOLI	Q47185	escherichia
8	5	29.4	72	1	HST3_ECOLI	P07965	escherichia
9	5	29.4	76	1	RS37_YEAST	P05759	saccharomyc
10	5	29.4	90	1	HQ37_HAEIN	P44437	haemophilus
11	5	29.4	90	1	Y41G_RHITSN	P55490	rhizobium s
12	5	29.4	95	1	RR20_GUTH	O78486	guillardia
13	5	29.4	98	1	SV13_HUMAN	Q99616	homo sapien
14	5	29.4	101	1	HFO_ECOLI	P25521	escherichia
15	5	29.4	101	1	HFO_SALTY	Q56059	salmonella
16	5	29.4	102	1	YNI6_YEAST	P48232	saccharomyc
17	5	29.4	113	1	KV2D_HUMAN	P01617	homo sapien
18	5	29.4	113	1	KV2E_MOUSE	P03976	mus musculus
19	5	29.4	113	1	KV2F_MOUSE	P01630	mus musculus
20	4	23.5	19	1	PHSL_DESEN	P13066	desulfotomac
21	4	23.5	20	1	DEF6_DERFA	P49276	dermatophag
22	4	23.5	20	1	YPRB_SERMA	P22581	serratia ma
23	4	23.5	24	1	CT3L_LITIC1	P81851	litorea cit
24	4	23.5	31	1	SODC_STRHE	P81163	striga herm
25	4	23.5	31	1	Y822_BORBU	O51762	borrellia bu
26	4	23.5	34	1	Y870_HAEIN	P44065	haemophilus
27	4	23.5	36	1	AMPL_PIG	P28839	sus scrofa
28	4	23.5	36	1	RL7_COXB	O87902	coxiella bu
29	4	23.5	43	1	TYBY_HUMAN	O14604	homo sapien
30	4	23.5	44	1	YX9_ODOSI	P49835	odonteella s
31	4	23.5	45	1	YOOD_BPT4	Q01434	bacteriophag
32	4	23.5	48	1	ATP8_CANPA	P17345	candida par
33	4	23.5	48	1	RL34_MYCPN	P78006	mycoplasma

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKKL 13  
|||||  
Db 14 SSKKL 19

RESULT 2  
DNBI\_BFDV  
ID DNBI\_BFDV STANDARD; PRT; 61 AA.  
AC P13893;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 01-JAN-1990 (Rel. 13, Last annotation update)  
DE DNA-BINDING PROTEIN (AGNOPROTEIN).  
OS Budgerigar fledgling disease virus (BFDV).  
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.  
OX NCBI\_TaxID=10625;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=88265888; PubMed=2838972;  
RX Rott O., Kroeger M., Mueller H., Hobom G.;  
RT "The genome of budgerigar fledgling disease virus, an avian  
polyomavirus";  
RL Virology 165:74-86(1988).  
CC -|- FUNCTION: THIS PROTEIN, ENCODED BY THE AGNOGENE, MAY HAVE A  
REGULATORY ROLE IN NUCLEIC ACID-PROTEIN INTERACTIONS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; M20775; AAB59756.1; -  
DR PIR; A29194; DNPVBF.  
KW DNA-binding.  
SQ SEQUENCE 61 AA; 6845 MW; 4B628909DA238673 CRC64;

Query Match 29.4%; Score 5; DB 1; Length 61;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHS 8  
|||||  
Db 12 SLLHS 16

RESULT 3  
YM45\_CAEEL  
ID YM45\_CAEEL STANDARD; PRT; 62 AA.  
AC F34521;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE HYPOTHETICAL 7.0 KDA PROTEIN K11H3.5 IN CHROMOSOME III.  
GN K11H3.5.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae.  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=BRISTOL N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,

RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans";  
RL Nature 368:32-38(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; Z22180; CAA80177.1; -  
DR PIR; S40758; S40758.  
DR WormPeP; K11H3.5; CE00265.  
KW Hypothetical protein.  
SQ SEQUENCE 62 AA; 7047 MW; 75FC9EC7916D8AB5 CRC64;

Query Match 29.4%; Score 5; DB 1; Length 62;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 EKLRR 15  
|||||  
Db 28 EKLRR 32

RESULT 4  
YORQ\_TTV1  
ID YORQ\_TTV1 STANDARD; PRT; 63 AA.  
AC P19301;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 01-NOV-1990 (Rel. 16, Last annotation update)  
DE HYPOTHETICAL 7.3 KDA PROTEIN.  
OS Thermoproteus tenax virus 1 (strain KRAL) (TTV1).  
OC Viruses; dsDNA viruses, no RNA stage; Lipothrixviridae;  
OC Lipothrixvirus.  
OX NCBI\_TaxID=10480;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Neumann H.;  
RL Submitted (MAR-1989) to the EMBL/GenBank/DBJ databases.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X14855; CAA32997.1; -  
DR Hypothetical protein.  
SQ SEQUENCE 63 AA; 7388 MW; 709617CD36C5A35B CRC64;

Query Match 29.4%; Score 5; DB 1; Length 63;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHS 8  
|||||  
Db 35 SLLHS 39

```

RESULT 5
RL29_PYRHO STANDARD; PRT; 68 AA.
AC Q74094;
DT 30-MAY-2000 (Rel. 39, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L29P.
GN RPL29P OR PHS048.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kostugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudo H. Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
CC
CC -!- SIMILARITY: BELONGS TO THE L29P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP000007; BAA30887.1; ALT_INIT.
DR InterPro; IPR001854; -.
DR Pfam; PF00831; Ribosomal_L29; 1.
DR PROSITE; PS00579; RIBOSOMAL_L29; 1.
KW Ribosomal protein.
SQ SEQUENCE 68 AA; 8038 MW; 06AD145F48E4CD23 CRC64;
Query Match 29.4%; Score 5; DB 1; Length 68;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 KEKLR 14
DB 59 KEKLR 63
|||||
-----
RESULT 6
HST1_ECOLI STANDARD; PRT; 72 AA.
AC P01559; Q47653;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HEAT-STABLE ENTEROTOXIN ST-1A/ST-P PRECURSOR.
GN STAL.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-Tn1681;
RX MEDLINE=81054703; PubMed=6254008;
RA So M., McCarthy B.J.;
RT "Nucleotide sequence of the bacterial transposon Tn1681 encoding a
RT heat-stable (ST) toxin and its identification in enterotoxigenic
RT Escherichia coli strains."

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 77:4011-4015(1980).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=18D;
RX MEDLINE=90368614; PubMed=2203756;
RA Dallas W.S.;
RT "The heat-stable toxin I gene from Escherichia coli 18D."
RL J. Bacteriol. 172:5490-5493(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85249571; PubMed=2990268;
RA Sekizaki T., Akashi H., Terakado N.;
RT "Nucleotide sequences of the genes for Escherichia coli heat-stable
RT enterotoxin I of bovine, avian, and porcine origins."
RL Am. J. Vet. Res. 46:909-912(1985).
RN [4]
RP DISULFIDE BONDS.
RX MEDLINE=87191003; PubMed=3552731;
RA Shimonishi Y., Hidaka Y., Koizumi M., Hane M., Aimoto S., Takeda T.,
RA Miwatani T., Takeda Y.;
RT "Mode of disulfide bond formation of a heat-stable enterotoxin (STH)
RT produced by a human strain of enterotoxigenic Escherichia coli."
RL FEBS Lett. 215:165-170(1987).
RN [5]
RP PROCESSING.
RX MEDLINE=90368584; PubMed=2203746;
RA Okamoto K., Takahara M.;
RT "Synthesis of Escherichia coli heat-stable enterotoxin Stp as a
RT pre-pro form and role of the pro sequence in secretion."
RL J. Bacteriol. 172:5260-5265(1990).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (0.89 ANGSTROMS) OF 59-71.
RX MEDLINE=94312375; PubMed=8038153;
RA Sato T., Ozaki H., Hata Y., Kitagawa Y., Katsube Y., Shimonishi Y.;
RT "Structural characteristics for biological activity of heat-stable
RT enterotoxin produced by enterotoxigenic Escherichia coli: X-ray
RT crystallography of weakly toxic and nontoxic analogs."
RL Biochemistry 33:8641-8650(1994).
CC -!- FUNCTION: TOXIN WHICH ACTIVATES THE PARTICULATE FORM OF GUANYLATE
CC CYCLASE AND INCREASES CYCLIC GMP LEVELS WITHIN THE HOST
CC INTESTINAL EPITHELIAL CELLS.
CC -!- DISEASE: BOTH HEAT-STABLE AND HEAT-LABILE ENTEROTOXINS ARE
CC PRODUCED BY PATHOGENIC STRAINS OF E.COLI AND AFFECT THE DIGESTIVE
CC TRACT OF MAMMALS.
CC -!- SIMILARITY: BELONGS TO THE HEAT-STABLE ENTEROTOXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V00612; CAA23883.1; -.
DR EMBL; M58746; AAA62776.1; -.
DR EMBL; M25607; AAA24653.1; -.
DR PIR; A01822; OHEC1.
DR PIR; J0374; J0374.
DR PIR; A36732; A36732.
DR PDB; 1ETL; 29-JAN-96.
DR PDB; 1ETN; 29-JAN-96.
DR InterPro; IPR001489; -.
DR Pfam; PF02048; Enterotoxin_HS; 1.
DR PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
KW Transposable element; Toxin; Enterotoxin; Signal; 3D-structure.
FT SIGNAL 1 19
FT PROPEP 20 54
FT PEPTIDE 55 72
FT DISULFID 59 64
FT DISULFID 60 68
FT DISULFID 63 71

```

FT CONFLICT 70 70 G -> P (IN REF. 3).  
SQ SEQUENCE 72 AA; 8075 MW; 92E8B76B3988264 CRC64;

Query Match 29.4%; Score 5; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKEK 12  
|||||  
Db 27 SSKEK 31

RESULT 7  
HST2\_ECOLI STANDARD; PRT; 72 AA.  
AC 047185;

DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE HEAT-STABLE ENTEROTOXIN A2 PRECURSOR (STA2).  
GN STA2.

OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=89108616; PubMed=2643580;  
RA Guzman-Verduzco L.M., Kupersztosch Y.M.;  
RT "Rectification of two Escherichia coli heat-stable enterotoxin allele sequences and lack of biological effect of changing the carboxy-terminal tyrosine to histidine.";  
RL Infect. Immun. 57:645-648(1989).  
CC -1- FUNCTION: TOXIN WHICH ACTIVATES THE PARTICULATE FORM OF GUANYLATE CYCLASE AND INCREASES CYCLIC GMP LEVELS WITHIN THE HOST  
CC INTESTINAL EPITHELIAL CELLS.  
CC -1- DISEASE: BOTH HEAT-STABLE AND HEAT-LABILE ENTEROTOXINS ARE  
CC PRODUCED BY PATHOGENIC STRAINS OF E.COLI AND EFFECT THE  
CC DIGESTIVE TRACT OF MAMMALS.  
CC -1- SIMILARITY: BELONGS TO THE HEAT-STABLE ENTEROTOXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M18345; AAA23729.1; -  
DR InterPro: IPR001489; -  
DR Pfam: PF02048; Enterotoxin\_HS; 1.  
DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.

KW Toxin; Enterotoxin; Signal.  
FT SIGNAL 1 19 POTENTIAL.  
FT PROPEP 20 53 BY SIMILARITY.  
FT PEPTIDE 54 72 HEAT-STABLE ENTEROTOXIN A2.  
FT DISULFID 59 64 BY SIMILARITY.  
FT DISULFID 60 68 BY SIMILARITY.  
FT DISULFID 63 71 BY SIMILARITY.  
SQ SEQUENCE 72 AA; 7895 MW; D87850306E06B260 CRC64;

Query Match 29.4%; Score 5; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKEK 12  
|||||  
Db 27 SSKEK 31

RESULT 8  
HST3\_ECOLI STANDARD; PRT; 72 AA.  
AC 007965; P26588;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE HEAT-STABLE ENTEROTOXIN A3/A4 PRECURSOR (STA3/STA4) (ST-IB) (ST-H).  
GN STA3 OR STA4.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89202548; PubMed=3071819;  
RA Stieglitz H., Cervantes L., Robledo R., Fonseca R., Covarrubias L.,  
RA Bolivar F., Kupersztosch Y.M.;  
RT "Cloning, sequencing, and expression in Ficol1-generated minicells of  
RT an Escherichia coli heat-stable enterotoxin gene.";  
RL Plasmid 20:42-53(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=153837-2;  
RX MEDLINE=83184648; PubMed=6341230;  
RA Moseley S.L., Hardy J.W., Huq M.I., Echeverria P., Falkow S.;  
RT "Isolation and nucleotide sequence determination of a gene encoding a  
RT heat-stable enterotoxin of Escherichia coli.";  
RL Infect. Immun. 39:1167-1174(1983).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90273381; PubMed=2190361;  
RA Zhou X., Shen L.P., Chi C.W.;  
RT "Isolation and nucleotide sequence determination of a gene encoding a  
RT heat-stable enterotoxin of Escherichia coli.";  
RL Toxicon 28:453-456(1990).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89108616; PubMed=2643580;  
RA Guzman-Verduzco L.M., Kupersztosch Y.M.;  
RT "Rectification of two Escherichia coli heat-stable enterotoxin allele  
RT sequences and lack of biological effect of changing the carboxy-  
RT terminal tyrosine to histidine.";  
RL Infect. Immun. 57:645-648(1989).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90034194; PubMed=2680769;  
RA Dwarakanath P., Visweswariah S.S., Subrahmanyam Y.V.B.K., Shanthi G.,  
RA Jagannatha H.M., Balganesht T.S.;  
RT "Cloning and hyperexpression of a gene encoding the heat-stable toxin  
RT of Escherichia coli.";  
RL Gene 81:219-226(1989).  
RN [6]  
RP SEQUENCE OF 54-72.  
RX MEDLINE=83105138; PubMed=6759126;  
RA Aimoto S., Takao T., Shimonishi Y., Hara S., Takeda T., Takeda Y.,  
RA Miwatani T.;  
RT "Amino-acid sequence of a heat-stable enterotoxin produced by human  
RT enterotoxigenic Escherichia coli.";  
RL Eur. J. Biochem. 129:257-263(1982).  
RN [7]  
RP DISULFIDE BONDS.  
RX MEDLINE=87191003; PubMed=3552731;  
RA Shimonishi Y., Hidaka Y., Koizumi M., Hane M., Aimoto S., Takeda T.,  
RA Miwatani T., Takeda Y.;  
RT "Mode of disulfide bond formation of a heat-stable enterotoxin (STh)  
RT produced by a human strain of enterotoxigenic Escherichia coli.";  
RL FEBS Lett. 215:165-170(1987).  
RN [8]  
RP PROCESSING.  
RX MEDLINE=90251166; PubMed=2187146;  
RA Rashed J.K., Guzman-Verduzco L.M., Kupersztosch Y.M.;  
RT "Two precursors of the heat-stable enterotoxin of Escherichia coli:

```

evidence of extracellular processing." ;
RL Mol. Microbiol. 4:265-273(1990).
CC -!- FUNCTION: TOXIN WHICH ACTIVATES THE PARTICULATE FORM OF GUANYLATE
CC CYCLASE AND INCREASES CYCLIC GMP LEVELS WITHIN THE HOST
CC INTESTINAL EPITHELIAL CELLS.
CC -!- DISEASE: BOTH HEAT-STABLE AND HEAT-LABILE ENTEROTOXINS ARE
CC PRODUCED BY PATHOGENIC STRAINS OF E.COLI AND EFFECT THE
CC DIGESTIVE TRACT OF MAMMALS.
CC -!- SIMILARITY: BELONGS TO THE HEAT-STABLE ENTEROTOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03311; AAA24652.1; -
DR EMBL; M34916; AAA23990.1; -
DR EMBL; M18346; AAA23730.1; -
DR EMBL; M29255; AAA24686.1; -
DR PIR; JS0292; OHECIB.
DR PIR; JT0373; OHECA.
DR HSP; P01559; IETN.
DR InterPro; IPR001489; -.
DR Pfam; PF02048; Enterotoxin_HS; 1.
DR PROSITE; PS00273; ENTEROTOXIN_H_STABLE; 1.
DR Toxin; Enterotoxin; Signal.
KW SIGNAL 1 19
KW FT PROPEP 20 53
KW FT PEPTIDE 54 72 HEAT-STABLE ENTEROTOXIN A3/A4.
KW FT DISULFID 59 64
KW FT DISULFID 60 68
KW FT DISULFID 63 71
KW FT CONFLICT 19 19 A -> P (IN REF. 2).
KW SEQUENCE 72 AA; 7909 MW; 1C5C9292BFCBA6BA CRC64;

Query Match 29.4%; Score 5; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]

Johnson D., Johnston L., Langston Y., Latreille P., Le T.,  
Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,  
Rifken L., Riles L., Taich A., Trevasakis E., Vignati D.,  
Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;  
Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE OF 1-19.  
Otake E., Higo K.-I., Itoh T.;  
"Yeast ribosomal proteins: VIII. Isolation of two proteins and  
sequence characterization of twenty-four proteins from cytoplasmic  
ribosomes."  
Mol. Gen. Genet. 195:544-546(1984).  
[4]  
IDENTIFICATION OF PROTEIN.  
MEDLINE=89181925; PubMed=2538753;  
Finley D., Bartel B., Varshavsky A.;  
"The tails of ubiquitin precursors are ribosomal proteins whose  
fusion to ubiquitin facilitates ribosome biogenesis."  
Nature 338:394-401(1989).  
-1- MISCELLANEOUS: THIS RIBOSOMAL PROTEIN IS SYNTHESIZED AS A  
C-TERMINAL EXTENSION PROTEIN (CEP) OF UBIQUITIN.  
-1- SIMILARITY: BELONGS TO THE S27AE FAMILY OF RIBOSOMAL PROTEINS.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X05730; CAA29197.1; ALT\_INIT.  
EMBL; U17246; AAB67466.1; ALT\_INIT.  
PIR; C29456; UOBYR7.  
SGD; S0004157; RPS31.  
InterPro; IPR002906; -  
Pfam; PF01599; Ribosomal\_S27; 1.  
Ribosomal protein; Zinc-finger; Metal-binding.  
DOMAIN 1 23  
ZN\_FING 45 68  
SEQUENCE 76 AA; 8677 MW; 691AF9DD013D0BFF CRC64;  
-----  
Query Match 29.4%; Score 5; DB 1; Length 76;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 12 KLRRE 16  
|||||  
Db 40 KLRRE 44  
  
RESULT 10  
HFQ\_HAEIN STANDARD; PRT; 90 AA.  
ID HFQ\_HAEIN  
AC P44437;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HOST FACTOR-I PROTEIN (HF-I).  
GN HFQ OR H10411.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
SEQUENCE FROM N.A.  
STRAIN=RD / KW20 / ATCC 51907;  
MEDLINE=95350630; PubMed=7542800;  
RA Fleisichmann R.B., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L., Glodek A., Kelley J.M.,  
Rosenberg M., White O., Adams M.D., Clayton R.A., Kirkness E.F.,<

RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 RT influenzae Rd.";  
 RL Science 269:496-512(1995).  
 CC -I- FUNCTION: REQUIRED FOR PHAGE Q BETA RNA-DIRECTED SYNTHESIS OF  
 CC COMPLEMENTARY MINUS-STRAND RNA (BY SIMILARITY).  
 CC -I- SUBUNIT: HOMOHXAMER (BY SIMILARITY).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U32724; AAC22070.1; -  
 CC TIGR: H10411; -  
 CC RNA-binding.  
 CC INIT\_MET 0 0 BY SIMILARITY.  
 CC SEQUENCE 90 AA: 10118 MW: 109104 ECF1F7F029 CRC64;  
 CC  
 CC Query Match 29.4%; Score 5; DB 1; Length 90;  
 CC Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 13 LRER 17  
 CC |  
 CC Db 14 LRER 18  
 CC  
 CC RESULT 11  
 CC Y4IG\_RHISN STANDARD; PRT; 90 AA.  
 CC ID Y4IG\_RHISN  
 CC AC P5490;  
 CC DT 01-NOV-1997 (Rel. 35, Created)  
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 CC DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 CC DE HYPOTHETICAL 10.5 KDA PROTEIN Y4IG.  
 CC GN Y4IG.  
 CC OS Rhizobium sp. (strain NGR234).  
 CC OG Plasmid sym PNGR234a.  
 CC OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 CC OC Rhizobiaceae; Rhizobium.  
 CC OX NCBI\_TaxID=394;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC RX MEDLINE=97305956; PubMed=9163424;  
 CC RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,  
 RA Ferret X.;  
 RT "Molecular basis of symbiosis between Rhizobium and legumes.";  
 RL Nature 387:394-401(1997).  
 CC -I- SIMILARITY: NONE OBVIOUS OUTSIDE OF SIMILARITY WITH FA2.  
 CC POTENTIAL FRAGMENT.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE000078; AAB91702.1; -  
 CC Hypothetical protein; Plasmid.  
 CC SIMILAR 1 74 HIGHLY SIMILAR TO FA2.  
 CC SEQUENCE 90 AA: 10485 MW: 1380048D6D5CC9EA CRC64;  
 CC

Query Match 29.4%; Score 5; DB 1; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 13 LRER 17  
 |  
 Db 46 LRER 50  
 RESULT 12  
 RR20\_GUTH STANDARD; PRT; 95 AA.  
 ID RR20\_GUTH  
 AC O78486;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S20.  
 GN RPS20.  
 OS Guillardia theta (Cryptomonas phi).  
 OG Chloroplast.  
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
 OX NCBI\_TaxID=55529;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99128221; PubMed=9929392;  
 RA Douglas S.E., Penny S.L.;  
 RT "The plastid genome of the cryptophyte alga, Guillardia theta:  
 RT complete sequence and conserved syntenic groups confirm its common  
 RT ancestry with red algae.";  
 RL J. Mol. Evol. 48:236-244(1999).  
 CC -I- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA  
 CC (BY SIMILARITY).  
 CC -I- SIMILARITY: BELONGS TO THE S20P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF041468; AAC35677.1; -  
 CC InterPro: IPR002583;  
 CC DR Pfam: PF01649; Ribosomal\_S20p; 1.  
 CC KW Ribosomal protein; rRNA-binding; Chloroplast.  
 CC SEQUENCE 95 AA: 10822 MW: 45384A5C6396C583 CRC64;  
 CC  
 CC Query Match 29.4%; Score 5; DB 1; Length 95;  
 CC Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 1 KKISL 5  
 CC |  
 CC Db 91 KKISL 95  
 CC  
 CC RESULT 13  
 CC SY13\_HUMAN STANDARD; PRT; 98 AA.  
 ID SY13\_HUMAN  
 AC Q99616;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE SMALL INDUCIBLE CYTOKINE A13 PRECURSOR (MONOCYTE CHEMOTACTIC PROTEIN  
 DE 4) (MCP-4) (MONOCYTE CHEMOATTRACTANT PROTEIN 4) (CK-BETA-10) (NCC-1).  
 GN SCYA13 OR MCP4 OR NCC1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 CC

```

CC -1- PTM: ONE MAJOR ISOFORM MCP-4, AND TWO MINOR ISOFORMS (LA)MCP-4 AND
CC (FN)MCP-4 ARE PRODUCED BY DIFFERENTIAL SIGNAL CLEAVAGE.
CC (LA)MCP-4 IS ABOUT 30 FOLD LESS ACTIVE THAN MCP-4.
CC -1- MASS SPECTROMETRY: MW=9314; MW_ERR=30; METHOD=MALDI; RANGE=17-98.
CC -1- MASS SPECTROMETRY: MW=8760; MW_ERR=30; METHOD=MALDI; RANGE=22-98.
CC -1- MASS SPECTROMETRY: MW=8575; MW_ERR=30; METHOD=MALDI; RANGE=24-98.
CC -1- MISCELLANEOUS: THIS PROTEIN CAN BIND HEPARIN.
CC -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE
CC C-C) (CHEMOKINE CC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U46767; AAB38703.1; -.
CC DR EMBL; AC002482; AAB67307.1; -.
CC DR EMBL; X98306; CAA66950.1; -.
CC DR EMBL; U59808; AAD09362.1; -.
CC DR EMBL; AJ001634; CAA04888.1; -.
CC DR HSP; P13500; IDOL.
CC DR MIM; 601391; -.
CC DR InterPro; IPR000827; -.
CC DR InterPro; IPR001811; -.
CC DR Pfam; PF00048; IL8; 1.
CC DR PROSITE; PS00472; SMALL_CYTOKINES_CC; 1.
CC DR Cytokine; Chemotaxis; Signal; Glycoprotein; Inflammatory response.
CC KW SIGNAL 1 23
CC FT CHAIN 24 98 SMALL INDUCIBLE CYTOKINE A13.
CC FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
CC FT DISULFID 34 58 BY SIMILARITY.
CC FT DISULFID 35 74 BY SIMILARITY.
CC FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 98 AA; 10986 MW; 612688DFCD308873 CRC64;
CC -----
Query Match 29.4%; Score 5; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKISL 5
DB 41 KKISL 45
|||||
-----
RESULT 14
HFQ_ECOLI STANDARD; PRT; 101 AA.
AC P25521;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HOST FACTOR-I PROTEIN (HF-I).
GN HFQ.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OC NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-44.
RX MEDLINE-9121185; PubMed-2020545;
RA Kajitani M., Ishihama A.
RT "Identification and sequence determination of the host factor gene
RT for bacteriophage Q beta."
RL Nucleic Acids Res. 19:1063-1066(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE-95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,

```

```

RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RL region from 92.8 through 100 minutes.";
RN Nucleic Acids Res. 23:2105-2119(1995).
RP [3]
RC SEQUENCE OF 88-101 FROM N.A.
RX STRAIN=K12;
RY MEDLINE=94068503; PubMed=8248183;
RA Noble J.A., Innis M.A., Koonin E.V., Rudd K.E., Banuett F.,
RA Herskowitz I.;
RT "The Escherichia coli hflA locus encodes a putative GTP-binding
RT protein and two membrane proteins, one of which contains a
RT protease-like domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:10866-10870(1993).
RN [4]
RP SEQUENCE OF 1-5.
RX MEDLINE=75151506; PubMed=805130;
RA Carmichael G.G., Weber K., Niveleau A., Wahba A.J.;
RT "The host factor required for RNA phage Qbeta RNA replication in
RT vitro. Intracellular location, quantitation, and purification by
RT polyacrylate-cellulose chromatography.";
RL J. Biol. Chem. 250:3607-3612(1975).
RN [5]
RP SEQUENCE OF 1-18, AND DNA-BINDING.
RX MEDLINE=97396137; PubMed=9245691;
RA Takada A., Wachi M., Kadow A., Takamura M., Nagai K.;
RT "DNA binding properties of the hfg gene product of Escherichia coli.";
RL Biochem. Biophys. Res. Commun. 236:576-579(1997).
CC -1- FUNCTION: REQUIRED FOR PHAGE Q BETA RNA-DIRECTED SYNTHESIS OF
CC COMPLEMENTARY MINUS-STRAND RNA.
CC LINEAR DNA AND THE BINDING OF SEEMS TO BE SEQUENCE-NONSPECIFIC.
CC -1- SUBUNIT: HOMOHEXAMER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00743; BAA00644.1; -
DR EMBL; U14003; AAA97088.1; -
DR EMBL; AE000489; AAC77129.1; -
DR EMBL; U00005; AAC43397.1; -
DR PIR; S26832; S26832.
DR EcoGene; EC10438; hfg
KW RNA-binding; DNA-binding.
FT INIT_MET 0
SQ SEQUENCE 101 AA; 11035 MW; 2B435A25D359C9DD CRC64;

Query Match 29.4%; Score 5; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRRER 17
DB 14 LRRER 18

RESULT 15
HFQ_SALTY STANDARD; PRT; 101 AA.
AC Q56059;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HOST FACTOR-I PROTEIN (HF-I).
GN HFQ
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.

```

```

OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RA Brown L., Elliott T.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR PHAGE Q BETA RNA-DIRECTED SYNTHESIS OF
CC COMPLEMENTARY MINUS-STRAND RNA.
CC -1- SUBUNIT: HOMOHEXAMER (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U48735; AAA99108.1; -
DR StyGene; SG10653; hfg.
KW RNA-binding.
FT INIT_MET 0
SQ SEQUENCE 101 AA; 11042 MW; 3A04C2E1EF090D8A CRC64;

Query Match 29.4%; Score 5; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRRER 17
DB 14 LRRER 18

Search completed: September 15, 2001, 12:52:33
Job time: 107 sec

```





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:48:00 ; Search time 12.82 Seconds  
(without alignments)  
101.011 Million cell updates/sec

Title: us-09-389-000-2\_COPY\_134\_150

Perfect score: 17

Sequence: 1 KKISLLHSSKEKLRER 17

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 39557

Minimum DB seq length: 0

Maximum DB seq length: 116

Post-processing: Listing first 45 summaries

Database : PIR.68.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	35.3	42	2 T07291	hypothetical prote
2	6	35.3	68	1 IABY3	proteinase A inhib
3	6	35.3	90	2 H69159	hypothetical prote
4	6	35.3	94	2 T12719	hypothetical prote
5	6	35.3	100	2 PH1075	Ig kappa chain V r
6	6	35.3	109	2 PH0089	Ig kappa chain V r
7	5	29.4	28	2 T14905	hypothetical prote
8	5	29.4	39	2 S75759	hypothetical prote
9	5	29.4	42	2 T07248	hypothetical prote
10	5	29.4	43	2 F81505	hypothetical prote
11	5	29.4	47	2 T07311	hypothetical prote
12	5	29.4	48	2 S61469	hypothetical prote
13	5	29.4	48	2 G82172	p83/100 protein -
14	5	29.4	53	2 G82813	p83/100 protein -
15	5	29.4	60	2 H71186	probable ribosomal
16	5	29.4	61	1 DNVPBF	DNA-binding protei
17	5	29.4	62	2 S40758	hypothetical prote
18	5	29.4	64	2 F70224	hypothetical prote
19	5	29.4	71	2 H64454	hypothetical prote
20	5	29.4	72	1 QHEC1	heat-stable entero
21	5	29.4	72	1 QHEC1B	heat-stable entero
22	5	29.4	72	1 QHEC4	heat-stable entero
23	5	29.4	77	2 JQ1295	hypothetical 8.8K
24	5	29.4	78	2 S61468	p83/100 protein -
25	5	29.4	78	2 C71919	hypothetical prote
26	5	29.4	78	2 S61473	p83/100 protein -
27	5	29.4	78	2 S61475	p83/100 protein -
28	5	29.4	78	2 S61474	p83/100 protein -
29	5	29.4	84	2 S61465	p83/100 protein -

30	5	29.4	84	2	S61466	p83/100 protein -
31	5	29.4	87	2	S34091	Ig kappa chain V r
32	5	29.4	87	2	C82334	probable host fact
33	5	29.4	87	2	C84494	hypothetical prote
34	5	29.4	87	2	A64577	hypothetical prote
35	5	29.4	87	2	T22931	hypothetical prote
36	5	29.4	88	2	S60837	M protein precursor
37	5	29.4	90	2	E84865	hypothetical prote
38	5	29.4	91	2	D64066	host factor I - Ha
39	5	29.4	92	2	A28762	Ig kappa chain V r
40	5	29.4	92	2	T03617	retrovirus-related
41	5	29.4	92	2	T03665	reverse transcript
42	5	29.4	92	2	G82848	host factor-I prot
43	5	29.4	93	2	D69997	hypothetical prote
44	5	29.4	94	2	T46349	hypothetical prote
45	5	29.4	96	2	S04824	U1 snRNP 70K prote

ALIGNMENTS

RESULT 1

T07291

hypothetical protein 42c - Chlorella vulgaris chloroplast

C:Species: chloroplast Chlorella vulgaris

C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000

C:Accession: T07291

R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Na

Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997

A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga C

A:Reference number: Z15985; MUID:97303241

A:Accession: T07291

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-42 <WAK>

A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57939.1; PID:g2224455

C:Genetics:

A:Genome: chloroplast

C:Keywords: chloroplast

Query Match 35.3%; Score 6; DB 2; Length 42;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKISLL 6

Db 34 KKISLL 39

RESULT 2

IABY3

Proteinase A inhibitor 3 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YM8010.04c; protein YMR174c

C:Species: Saccharomyces cerevisiae

C:Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 21-Jul-2000

C:Accession: A01334; S16692; S55121

R:Biedermann, K.; Montali, U.; Martin, B.; Svendsen, I.; Ottesen, M.

Carlsberg Res. Commun. 45, 225-235, 1980

A:Title: The amino acid sequence of proteinase A inhibitor 3 from baker's yeast.

A:Reference number: A01334

A:Accession: A01334

A:Molecule type: protein

A:Residues: 1-68 <BI>

A>Note: nearly all of the inhibitory activity is present in the peptide consisting of

R:Schu, P.; Wolf, D.H.

FEBS Lett. 283, 78-84, 1991

A:Title: The proteinase yscA-inhibitor, I(A)(3), gene. Studies of cytoplasmic protein

A:Reference number: S16692; MUID:91243884

A:Accession: S16692

A:Molecule type: DNA

A:Residues: 1-68 <SCH>

A:Cross-references: EMBL:X60050; NID:g4094; PIDN:CAA42650.1; PID:g4095

R; Churcher, C.M.  
Submitted to the EMBL Data Library, June 1995  
A: Reference number: S55118  
A: Accession: S55121  
A: Molecule type: DNA  
A: Residues: 1-68 <CHU>  
A: Cross-references: EMBL:Z49808; NID:g8544440; PIDN:CAA89907.1; PID:g8544444; GSPDB:GN0001  
A: Experimental source: strain AB972  
C: Genes: SGD:PAI3; MIPS:YMR174C  
A: Gene: SGD:PAI3; MIPS:YMR174C  
A: Cross-references: SGD:S0004786; MIPS:YMR174C  
A: Map position: 13R  
C: Superfamily: proteinase A inhibitor 3  
C: Keywords: acetylated amino end; blocked amino end; proteinase inhibitor  
F: 1/Modified site: acetylated amino end (Met) #status experimental

Query Match 35.3%; Score 6; DB 1; Length 68;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKEKL 13  
Db 14 SSKEKL 19  
|||||

RESULT 3  
H69159  
hypothetical protein MTH458 - Methanobacterium thermoautotrophicum (strain Delta H)  
C: Species: Methanobacterium thermoautotrophicum  
C: Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C: Accession: H69159  
R: Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.;  
Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A: Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func  
A: Reference number: A69000; MUID:198037514  
A: Accession: H69159  
A: Status: preliminary; nucleic acid sequence not shown; translation not shown  
A: Molecule type: DNA  
A: Residues: 1-90 <MTH>  
A: Cross-references: GB:AE000830; GB:AE000666; NID:g2621523; PIDN:AAB84964.1; PID:g262152  
A: Experimental source: strain Delta H  
C: Genes: MTH458

Query Match 35.3%; Score 6; DB 2; Length 90;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KEKLRR 15  
Db 36 KEKLRR 41  
|||||

RESULT 4  
T12719  
hypothetical protein 3 - Methanobacterium phage psiM2  
C: Species: Methanobacterium phage psiM2  
C: Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 24-Sep-1999  
C: Accession: T12719  
R: Pfister, P.; Wasserfallen, A.; Stettler, R.; Leisinger, T.  
submitted to the EMBL Data Library, May 1998  
A: Description: Archaeophage PsiM2 complete genomic DNA.  
A: Reference number: Z17578  
A: Accession: T12719  
A: Status: translated from GB/EMBL/DBJ  
A: Molecule type: DNA  
A: Residues: 1-94 <PFI>  
A: Cross-references: EMBL:AF065411; NID:g3249585; PID:g3249588; PIDN:AAC27042.1  
A: Experimental source: host Methanobacterium thermoautotrophicum strain Marburg

Query Match 35.3%; Score 6; DB 2; Length 94;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 SKEKLR 14  
Db 72 SKEKLR 77  
|||||

RESULT 5  
PH1075  
Ig light chain V region (clone 17s.128) - mouse (fragment)  
C: Species: Mus musculus (house mouse)  
C: Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jan-2000  
C: Accession: PH1075  
J: Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.  
J. Exp. Med. 176, 761-779, 1992  
A: Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective  
A: Reference number: PH0971; MUID:92381444  
A: Accession: PH1075  
A: Status: nucleic acid sequence not shown  
A: Molecule type: mRNA  
A: Residues: 1-100 <TIL>  
A: Experimental source: B cell, strain [NZB x NZW]F1  
C: Superfamily: immunoglobulin V region; immunoglobulin homology  
C: Keywords: immunoglobulin  
F: 16-92/Domain: immunoglobulin homology <IMM>

Query Match 35.3%; Score 6; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHSS 9  
Db 28 SLLHSS 33  
|||||

RESULT 6  
PH0089  
Ig kappa chain V region (anti-cyclosporin A and B) - mouse (fragment)  
C: Species: Mus musculus (house mouse)  
C: Date: 15-Jan-1993 #sequence\_revision 15-Jan-1993 #text\_change 21-Jan-2000  
C: Accession: PH0089  
R: Schmitter, D.; Poch, O.; Zeder, G.; Heinrich, G.F.; Kocher, H.P.; Quesniaux, V.F.J.  
Mol. Immunol. 27, 1029-1038, 1990  
A: Title: Analysis of the structural diversity of monoclonal antibodies to cyclosporin  
A: Reference number: PH0087; MUID:91042649  
A: Accession: PH0089  
A: Molecule type: mRNA  
A: Residues: 1-109 <SCH>  
C: Superfamily: immunoglobulin V region; immunoglobulin homology  
C: Keywords: heterotetramer; immunoglobulin  
F: 16-95/Domain: immunoglobulin homology <IMM>  
F: 24-39/Region: complementarity-determining 1  
F: 55-61/Region: complementarity-determining 2  
F: 94-101/Region: complementarity-determining 3

Query Match 35.3%; Score 6; DB 2; Length 109;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHSS 9  
Db 28 SLLHSS 33  
|||||

RESULT 7  
T14905  
hypothetical protein - parsley

C:Species: Petroselinum crispum (parsley)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T14905  
R:Feldbrugge, M.; Sprenger, M.; Dinkelbach, M.; Yazaki, K.; Harter, K.; Weisshaar, B.  
Plant Cell 6, 1607-1621, 1994  
A:Title: Functional analysis of a light-responsive plant bZIP transcriptional regulator.  
A:Reference number: Z18259; MUID:95128172  
A:Accession: T14905  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-28 <FEL>  
A:Cross-references: EMBL:S75395; NID:g913201; PID:e194726

Query Match 29.4%; Score 5; DB 2; Length 28;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLLHS 8  
|||||  
Db 13 SLLHS 17

RESULT 8  
S75759  
hypothetical protein sgl0001 - Synecocystis sp. (strain PCC 6803)  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S75759  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis  
S.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S75759  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-39 <KAN>  
A:Cross-references: EMBL:D64003; GB:AB001339; NID:gl001200; PIDN:BAAL0494.1; PID:d101114  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Start codon: GTG

Query Match 29.4%; Score 5; DB 2; Length 39;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KISLL 6  
|||||  
Db 9 KISLL 13

RESULT 9  
T07248  
hypothetical protein 42a - Chlorella vulgaris chloroplast  
C:Species: Chloroplast Chlorella vulgaris  
C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000  
C:Accession: T07248  
R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Tsudzuki, J.; Nakas  
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlo  
A:Reference number: Z15985; MUID:97303241  
A:Accession: T07248  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-42 <WAK>  
A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57895.1; PID:g2224411  
C:Genetics:  
A:Genome: chloroplast  
C:Keywords: chloroplast

Query Match 29.4%; Score 5; DB 2; Length 47;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSKEK 12  
|||||  
Db 3 SSKEK 7

RESULT 12  
S61469  
p83/100 protein - Borrelia afzelii (strain pLe and others) (fragment)

Query Match 29.4%; Score 5; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 SKEKL 13  
|||||  
Db 19 SKEKL 23

RESULT 10  
F81505  
hypothetical protein CP1078 [imported] - Chlamydomonada pneumoniae (strain AR39)  
C:Species: Chlamydomonada pneumoniae, Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
C:Accession: F81505  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
, C.; Dodson, R.; Gwinn, W.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
A:Reference number: A81500; MUID:20150255  
A:Accession: F81505  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-43 <REA>  
A:Cross-references: GB:AB002264; GB:AB002161; NID:g7189984; PIDN:AAF38850.1; PID:g718  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP1078

Query Match 29.4%; Score 5; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKISL 5  
|||||  
Db 7 KKISL 11

RESULT 11  
T07311  
hypothetical protein 47b - Chlorella vulgaris chloroplast  
C:Species: chloroplast Chlorella vulgaris  
C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000  
C:Accession: T07311  
R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Na  
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga C  
A:Reference number: Z15985; MUID:97303241  
A:Accession: T07311  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-47 <WAK>  
A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57959.1; PID:g2224475  
C:Genetics:  
A:Genome: chloroplast  
C:Keywords: chloroplast

Query Match 29.4%; Score 5; DB 2; Length 47;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSKEK 12  
|||||  
Db 3 SSKEK 7

RESULT 12  
S61469  
p83/100 protein - Borrelia afzelii (strain pLe and others) (fragment)

C:Species: Borrelia afzelii  
 A:Variety: strain p1ue and others  
 C:Date: 19-Mar-1997 #sequence\_revision 13-Mar-1998 #text\_change 08-Oct-1999  
 C:Accession: S61469; S61470; S61471  
 R:Roessler, D.; Eifert, H.; Jauris-Heipke, S.; Lehnert, G.; Preac-Mursic, V.; Teepe, J.  
 Med. Microbiol. Immunol. 184, 23-32, 1995  
 A:Title: Molecular and immunological characterization of the p83/100 protein of various  
 A:Reference number: S61461; MUID:96149106  
 A:Accession: S61469  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-48 <ROE>  
 A:Cross-references: EMBL:X81529; NID:g928973; PIDN:CAA57248.1; PID:g928974; EMBL:X81530;  
 A:Experimental source: strain p1ue; strain PKJ7; strain p9au  
 C:Keywords: surface antigen

Query Match 29.4%; Score 5; DB 2; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKEK 12  
 |||||  
 DB 27 SSKEK 31

RESULT 13  
 S61472  
 p83/100 protein - Borrelia afzelii (strain PWUDI) (fragment)  
 C:Species: Borrelia afzelii  
 A:Variety: strain PWUDI  
 C:Date: 19-Mar-1997 #sequence\_revision 05-Dec-1998 #text\_change 08-Oct-1999  
 C:Accession: S61472; S72307  
 R:Roessler, D.; Eifert, H.; Jauris-Heipke, S.; Lehnert, G.; Preac-Mursic, V.; Teepe, J.  
 Med. Microbiol. Immunol. 184, 23-32, 1995  
 A:Title: Molecular and immunological characterization of the p83/100 protein of various  
 A:Reference number: S61461; MUID:96149106  
 A:Accession: S61472  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-48 <ROE>  
 A:Cross-references: EMBL:X81535  
 A:Experimental source: strain PWUDI  
 R:Roessler, D.  
 submitted to the EMBL Data Library, September 1994  
 A:Reference number: S72307  
 A:Accession: S72307  
 A:Molecule type: DNA  
 A:Residues: 1-7,'F',8-48 <ROW>  
 A:Cross-references: EMBL:X81535; NID:g928977; PIDN:CAA57254.1; PID:g928978  
 A:Experimental source: strain PWUDI  
 C:Keywords: surface antigen

Query Match 29.4%; Score 5; DB 2; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKEK 12  
 |||||  
 DB 27 SSKEK 31

RESULT 14  
 G82813  
 hypothetical protein XF0376 [imported] - Xylella fastidiosa (strain 9a5c)  
 C:Species: Xylella fastidiosa  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C:Accession: G82813  
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
 Nature 406, 151-157, 2000  
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A:Reference number: A82515; MUID:20365717

A:Note: for a complete list of authors see reference number A59328 below  
 A:Accession: G82813  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-53 <SIM>  
 A:Cross-references: GB:AE003889; GB:AE003849; NID:g9105205; PIDN:AAF83186.1; GSPDB:GN  
 A:Experimental source: strain 9a5c  
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer  
 as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins  
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.  
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa  
 M.; Tshukako, M.H.; da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
 A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: XF0376

Query Match 29.4%; Score 5; DB 2; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 EKLRR 15  
 |||||  
 DB 14 EKLRR 18

RESULT 15  
 H71186  
 probable ribosomal protein L29 - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
 C:Accession: H71186  
 R:Kawabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se  
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu  
 DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil  
 A:Reference number: A71000; MUID:98344137  
 A:Accession: H71186  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-60 <KAW>  
 A:Cross-references: GB:AP000007; NID:g3236134; PIDN:BA030887.1; PID:g3258204  
 A:Experimental source: strain OH3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBa  
 C:Genetics:  
 A:Gene: PHS048

Query Match 29.4%; Score 5; DB 2; Length 60;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KEKLR 14  
 |||||  
 DB 51 KEKLR 55

Search completed: September 15, 2001, 12:51:30  
 Job time: 210 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:49:01 ; Search time 12.28 Seconds  
(without alignments)  
28.505 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_150  
Perfect score: 17  
Sequence: 1 KKTSLHSSKEKLRR 17

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 156483

Minimum DB seq length: 0  
Maximum DB seq length: 116

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgnl\_7/ptodata/1/iaa/5A\_COMB.pap.\*  
2: /cgnl\_7/ptodata/1/iaa/5B\_COMB.pap.\*  
3: /cgnl\_7/ptodata/1/iaa/6A\_COMB.pap.\*  
4: /cgnl\_7/ptodata/1/iaa/6B\_COMB.pap.\*  
5: /cgnl\_7/ptodata/1/iaa/PCTUS\_COMB.pap.\*  
6: /cgnl\_7/ptodata/1/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	41.2	52	US-08-630-916A-76	Sequence 76, Appl
2	6	35.3	112	US-08-477-877B-87	Sequence 87, Appl
3	6	35.3	112	US-08-477-877B-88	Sequence 88, Appl
4	6	35.3	112	US-08-472-281A-87	Sequence 87, Appl
5	6	35.3	112	US-08-472-281A-88	Sequence 88, Appl
6	6	35.3	112	US-08-477-989B-87	Sequence 87, Appl
7	6	35.3	112	US-08-477-989B-88	Sequence 88, Appl
8	6	35.3	112	US-09-240-214-30	Sequence 30, Appl
9	5	29.4	6	PCT-US93-12679-5	Sequence 5, Appl
10	5	29.4	16	US-08-438-123-1	Sequence 1, Appl
11	5	29.4	18	US-08-438-123-9	Sequence 9, Appl
12	5	29.4	21	US-08-704-170-28	Sequence 28, Appl
13	5	29.4	21	US-08-383-753-70	Sequence 70, Appl
14	5	29.4	21	US-08-586-772-70	Sequence 70, Appl
15	5	29.4	21	US-08-959-512-70	Sequence 70, Appl
16	5	29.4	21	PCT-US94-02631-28	Sequence 28, Appl
17	5	29.4	26	US-08-620-151-114	Sequence 114, Appl
18	5	29.4	38	US-08-176-500-65	Sequence 65, Appl
19	5	29.4	38	US-08-471-052A-65	Sequence 65, Appl
20	5	29.4	38	US-08-189-331-65	Sequence 65, Appl
21	5	29.4	38	US-08-471-939-65	Sequence 65, Appl
22	5	29.4	38	US-08-471-800-65	Sequence 65, Appl
23	5	29.4	38	US-08-471-068-65	Sequence 65, Appl
24	5	29.4	40	US-08-190-802A-197	Sequence 197, Appl
25	5	29.4	53	US-08-630-916A-83	Sequence 83, Appl
26	5	29.4	78	US-08-343-443B-10	Sequence 10, Appl
27	5	29.4	79	US-08-611-510-6	Sequence 6, Appl

28	5	29.4	85	1	US-07-667-276A-9	Sequence 9, Appl
29	5	29.4	94	1	US-08-519-777-31	Sequence 31, Appl
30	5	29.4	94	1	US-08-742-035-31	Sequence 31, Appl
31	5	29.4	94	2	US-08-777-019-31	Sequence 31, Appl
32	5	29.4	94	2	US-08-777-143-31	Sequence 31, Appl
33	5	29.4	94	3	US-08-775-414-31	Sequence 31, Appl
34	5	29.4	94	4	US-08-931-858E-31	Sequence 31, Appl
35	5	29.4	94	4	US-08-981-739-31	Sequence 31, Appl
36	5	29.4	95	4	US-08-931-858E-173	Sequence 173, Appl
37	5	29.4	95	4	US-08-981-739-173	Sequence 173, Appl
38	5	29.4	98	4	US-08-613-822-4	Sequence 4, Appl
39	5	29.4	102	1	US-08-519-777-1	Sequence 1, Appl
40	5	29.4	102	1	US-08-742-035-1	Sequence 1, Appl
41	5	29.4	102	2	US-08-777-019-1	Sequence 1, Appl
42	5	29.4	102	2	US-08-777-143-1	Sequence 1, Appl
43	5	29.4	102	3	US-08-478-087A-18	Sequence 18, Appl
44	5	29.4	102	3	US-09-106-486-1	Sequence 1, Appl
45	5	29.4	102	3	US-08-775-414-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-08-630-916A-76  
; Sequence 76, Application US/08630916A  
; Patent No. 6011137  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; APPLICANT: Kay, Brian K.  
; APPLICANT: Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/630,916A  
; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 896-8864/9741  
; INFORMATION FOR SEQ ID NO: 76:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 52 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-630-916A-76

Query Match 41.2%; Score 7; DB 3; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 SSKELR 14  
|||||||

Db 30 SSKELR 36

## RESULT 2

US-08-477-877B-87  
; Sequence 87, Application US/08477877B  
; Patent No. 5730979

## GENERAL INFORMATION:

; APPLICANT: Bazin, Herv  
; APPLICANT: Latinne, Dominique  
; TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation  
; NUMBER OF SEQUENCES: 96  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
; ADDRESSEE: Cecchi, Stewart & Olstein

; CITY: Roseland

; STATE: New Jersey

; COUNTRY: U.S.A.

; ZIP: 07068

## COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch diskette

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,877B

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/407,009

; FILING DATE: 29-MAR-1995

; APPLICATION NUMBER: 08/119,032

; FILING DATE: 09-SEP-1993

; APPLICATION NUMBER: 08/027,008

; FILING DATE: 05-MAR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Olstein, Elliot M.

; REGISTRATION NUMBER: 24,025

; REFERENCE/DOCKET NUMBER: 61750-146

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 87:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 112 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; MOLECULE TYPE: polypeptide

; FEATURE:

; NAME/KEY: Rat LO-CD2a light chain variable region.

US-08-477-877B-87

## Query Match

Best Local Similarity 35.3%; Score 6; DB 1; Length 112;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLLHSS 9

|||||

Db 28 SLLHSS 33

## RESULT 3

US-08-477-877B-88  
; Sequence 88, Application US/08477877B  
; Patent No. 5730979

## GENERAL INFORMATION:

; APPLICANT: Bazin, Herv

; APPLICANT: Latinne, Dominique

; TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation

; NUMBER OF SEQUENCES: 96

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
; ADDRESSEE: Cecchi, Stewart & Olstein  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: U.S.A.  
; ZIP: 07068

## COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch diskette

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,877B

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/407,009

; FILING DATE: 29-MAR-1995

; APPLICATION NUMBER: 08/119,032

; FILING DATE: 09-SEP-1993

; APPLICATION NUMBER: 08/027,008

; FILING DATE: 05-MAR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Olstein, Elliot M.

; REGISTRATION NUMBER: 24,025

; REFERENCE/DOCKET NUMBER: 61750-146

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 88:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 112 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; MOLECULE TYPE: polypeptide

; FEATURE:

; NAME/KEY: Humanized LO-CD2a light chain variable region.

US-08-477-877B-88

## Query Match

Best Local Similarity 35.3%; Score 6; DB 1; Length 112;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLLHSS 9

|||||

Db 28 SLLHSS 33

## RESULT 4

US-08-472-281A-87

; Sequence 87, Application US/08472281A

; Patent No. 5817311

; GENERAL INFORMATION:

; APPLICANT: Bazin, Herv

; APPLICANT: Latinne, Dominique

; TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Act

; NUMBER OF SEQUENCES: 96

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,

; ADDRESSEE: Cecchi, Stewart & Olstein

; STREET: 6 Becker Farm Road

; CITY: Roseland

; STATE: New Jersey

; COUNTRY: U.S.A.

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch diskette

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

```
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,281A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,009
; FILING DATE: 29-MAR-1995
; APPLICATION NUMBER: 08/119,032
; FILING DATE: 09-SEP-1993
; APPLICATION NUMBER: 08/027,008
; FILING DATE: 05-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 61750-142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; FEATURE:
; NAME/KEY: Rat LO-CD2a light chain variable region.
US-08-472-281A-87

Query Match 35.3%; Score 6; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLLHSS 9
Db 28 SLLHSS 33

RESULT 5
US-08-472-281A-88
; Sequence 88, Application US/08472281A
; Patent No. 5817311
; GENERAL INFORMATION:
; APPLICANT: Bazin, Herv
; APPLICANT: Latinne, Dominique
; TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,281A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,009
; FILING DATE: 29-MAR-1995
; APPLICATION NUMBER: 08/119,032
; FILING DATE: 09-SEP-1993
; APPLICATION NUMBER: 08/027,008
; FILING DATE: 05-MAR-1993
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 61750-142
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; FEATURE:
; NAME/KEY: Humanized LO-CD2a light chain variable region.
US-08-472-281A-88

Query Match 35.3%; Score 6; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLLHSS 9
Db 28 SLLHSS 33

RESULT 6
US-08-477-989B-87
; Sequence 87, Application US/08477989B
; Patent No. 5951983
; GENERAL INFORMATION:
; APPLICANT: Bazin, Herv
; APPLICANT: Latinne, Dominique
; APPLICANT: Kaplan, Ruth
; APPLICANT: Kieber-Emmons, Thomas
; APPLICANT: Postema, Christina E.
; APPLICANT: White-Scharf, Mary
; TITLE OF INVENTION: LO-CD2a Antibody and Uses Thereof for Inhibiting T-Cell Activation
; TITLE OF INVENTION: T-Cell Activation and Proliferation
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESSEE: Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,989B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,009
; FILING DATE: 29-MAR-1995
; APPLICATION NUMBER: 08/119,032
; FILING DATE: 09-SEP-1993
; APPLICATION NUMBER: 08/027,008
; FILING DATE: 05-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 61750-147
; TELECOMMUNICATION INFORMATION:
```

; TELEPHONE: 201-994-1700  
 ; TELEFAX: 201-994-1744  
 ; INFORMATION FOR SEQ ID NO: 87:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 112 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: polypeptide  
 ; FEATURE:  
 ; NAME/KEY: Rat LO-CD2a light chain variable  
 ; NAME/KEY: region.  
 ; US-08-477-989B-87

Query Match 35.3%; Score 6; DB 2; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHSS 9  
 Db 28 SLLHSS 33

## RESULT 7

; US-08-477-989B-88  
 ; Sequence 88, Application US/08477989B  
 ; Patent No. 5951983

; GENERAL INFORMATION:  
 ; APPLICANT: Bazin, Herv  
 ; APPLICANT: Latinne, Dominique  
 ; APPLICANT: Kaplan, Ruth  
 ; APPLICANT: Kieber-Emmons, Thomas  
 ; APPLICANT: Postema, Christina E.  
 ; APPLICANT: White-Scharf, Mary  
 ; TITLE OF INVENTION: LO-CD2a Antibody and Uses  
 ; TITLE OF INVENTION: Thereof for Inhibiting  
 ; TITLE OF INVENTION: T-Cell Activation and  
 ; NUMBER OF SEQUENCES: 96  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,  
 ; ADDRESSEE: Cecchi, Stewart & Olstein  
 ; STREET: 6 Becker Farm Road  
 ; CITY: Roseland  
 ; STATE: New Jersey  
 ; COUNTRY: U.S.A.  
 ; ZIP: 07068

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 inch diskette  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: MS-DOS  
 ; SOFTWARE: Wordperfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/477,989B  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/407,009  
 ; FILING DATE: 29-MAR-1995  
 ; APPLICATION NUMBER: 08/119,032  
 ; FILING DATE: 09-SEP-1993  
 ; APPLICATION NUMBER: 08/027,008  
 ; FILING DATE: 05-MAR-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Olstein, Elliot M.  
 ; REGISTRATION NUMBER: 24,025  
 ; REFERENCE/DOCKET NUMBER: 61750-147  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 201-994-1700  
 ; TELEFAX: 201-994-1744  
 ; INFORMATION FOR SEQ ID NO: 88:  
 ; SEQUENCE CHARACTERISTICS:

; LENGTH: 112 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: polypeptide  
 ; FEATURE:  
 ; NAME/KEY: Humanized LO-CD2a light chain  
 ; NAME/KEY: variable region.  
 ; US-08-477-989B-88

Query Match 35.3%; Score 6; DB 2; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHSS 9  
 Db 28 SLLHSS 33

## RESULT 8

; US-09-240-274-30  
 ; Sequence 30, Application US/09240274  
 ; Patent No. 6255455

; GENERAL INFORMATION:  
 ; APPLICANT: Siegel, Donald L.  
 ; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL  
 ; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF  
 ; FILE REFERENCE: 09596-4202

; CURRENT APPLICATION NUMBER: US/09/240,274  
 ; CURRENT FILING DATE: 1999-01-29  
 ; EARLIER APPLICATION NUMBER: 60/081,380  
 ; EARLIER FILING DATE: 1998-04-10  
 ; EARLIER APPLICATION NUMBER: 60/028,550  
 ; EARLIER FILING DATE: 1996-10-11  
 ; NUMBER OF SEQ ID NOS: 224  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 30  
 ; LENGTH: 112  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: anti-Rh(D) chain G01  
 ; US-09-240-274-30

Query Match 35.3%; Score 6; DB 4; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHSS 9  
 Db 27 SLLHSS 32

## RESULT 9

; PCT-US93-12679-5

; Sequence 5, Application PC/TUS9312679  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rath, Matthias  
 ; TITLE OF INVENTION: SYNTHETIC OLIGOPEPTIDES ANALOGOUS TO  
 ; TITLE OF INVENTION: PROTEIN SIGNAL SEQUENCES -METHODS OF IDENTIFICATION AND  
 ; TITLE OF INVENTION: METHODS OF USE  
 ; NUMBER OF SEQUENCES: 41  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SHELDON & MAK  
 ; STREET: 401 Florence Street  
 ; CITY: Palo Alto  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94301  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/12679  
FILING DATE: 30-DEC-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cranfill, Raymond B  
REGISTRATION NUMBER: 32,845  
REFERENCE/DOCKET NUMBER: RATH-10016PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-322-5333  
TELEFAX: 415-322-5499  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US93-12679-5

Query Match 29.4%; Score 5; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 8 SSKK 12  
DB 2 SSKK 6

RESULT 10  
US-08-438-123-1  
; Sequence 1, Application US/08438123  
; Patent No. 5552293  
; GENERAL INFORMATION:  
; APPLICANT: Lindholm et al  
; TITLE OF INVENTION: TUMOR ANTIGEN SPECIFIC ANTIBODY  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lowe, Price, LeBlanc & Becker  
; STREET: Suite 300, 99 Canal Center Plaza  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22314  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: DOS Text File  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/438,123  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/906,350  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: J.G. Mullins  
; REGISTRATION NUMBER: 33073  
; REFERENCE/DOCKET NUMBER: 149-011  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703 684 1111  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Polypeptide  
US-08-438-123-1

Query Match 29.4%; Score 5; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 4 SLLHS 8  
DB 5 SLLHS 9

RESULT 11  
US-08-438-123-9  
; Sequence 9, Application US/08438123  
; Patent No. 5552293  
; GENERAL INFORMATION:  
; APPLICANT: Lindholm et al  
; TITLE OF INVENTION: TUMOR ANTIGEN SPECIFIC ANTIBODY  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lowe, Price, LeBlanc & Becker  
; STREET: Suite 300, 99 Canal Center Plaza  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22314  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: DOS Text File  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/438,123  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/906,350  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: J.G. Mullins  
; REGISTRATION NUMBER: 33073  
; REFERENCE/DOCKET NUMBER: 149-011  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703 684 1111  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Polypeptide  
US-08-438-123-9

Query Match 29.4%; Score 5; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 5; Conservative 0; Mismatches 0; Indels 0;

QY 4 SLLHS 8  
DB 5 SLLHS 9

RESULT 12  
US-08-704-170-28  
; Sequence 28, Application US/08704170  
; Patent No. 5707626  
; GENERAL INFORMATION:  
; APPLICANT: Douvas, Angeline  
; APPLICANT: Takehana, Yoshi  
; APPLICANT: Ehresmann, Glenn  
; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
; IMMUNOINFECTION: IMMUNOINFECTION CLUSTER VIRUS INFECTIONS

NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Robbins, Berlinger & Carson  
STREET: 201 No. 570626th Figueroa Street, Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/704,170  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-704-170-28

Query Match 29.4%; Score 5; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 KLRRE 16  
Db 2 KLRRE 6

RESULT 13  
US-08-383-753-70  
Sequence 70, Application US/08383753  
Patent No. 5723584  
GENERAL INFORMATION:  
APPLICANT: Schatz, Peter J.  
TITLE OF INVENTION: Biotinylation of Proteins  
NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Khourie and Crew  
STREET: One Market Plaza, Steuart Tower  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/383,753  
FILING DATE: 03-FEB-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/099,991  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:

NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 1038.1  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-383-753-70

Query Match 29.4%; Score 5; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHS 8  
Db 5 SLLHS 9

RESULT 14  
US-08-586-772-70  
Sequence 70, Application US/08586772  
Patent No. 5874239  
GENERAL INFORMATION:  
APPLICANT: Schatz, Peter J.  
TITLE OF INVENTION: Biotinylation of Proteins  
NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Khourie and Crew  
STREET: One Market Plaza, Steuart Tower  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/586,772  
FILING DATE: 03-FEB-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/099,991  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 1038.1  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-586-772-70

Query Match 29.4%; Score 5; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLLHS 8  
|||||  
Db 5 SLLHS 9

## RESULT 15

US-08-959-512-70  
; Sequence 70, Application US/08959512  
; Patent No. 5932433  
; GENERAL INFORMATION:  
; APPLICANT: Schatz, Peter J.  
; TITLE OF INVENTION: Biotinylation of Proteins  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Hourie and Crew  
; STREET: One Market Plaza, Steuart Tower  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/959,512  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/383,753  
; FILING DATE: 03-FEB-1995  
; APPLICATION NUMBER: US 08/099,991  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 1038.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 70:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-959-512-70

Query Match 29.4%; Score 5; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLLHS 8  
|||||  
Db 5 SLLHS 9

Search completed: September 15, 2001, 12:51:50  
Job time: 169 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:47:20 ; Search time 17.45 Seconds  
(without alignments)  
59.061 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_150  
Perfect score: 17  
Sequence: 1 KKTSLHSSKEKLRRER 17

Scoring table: OLIGO  
Gapox 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 266123

Minimum DB seq length: 0

Maximum DB seq length: 116

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_0601.\*

- 1: /cgnl\_9/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /cgnl\_9/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /cgnl\_9/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /cgnl\_9/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /cgnl\_9/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /cgnl\_9/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /cgnl\_9/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /cgnl\_9/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /cgnl\_9/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /cgnl\_9/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /cgnl\_9/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /cgnl\_9/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /cgnl\_9/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /cgnl\_9/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /cgnl\_9/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /cgnl\_9/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /cgnl\_9/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /cgnl\_9/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /cgnl\_9/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /cgnl\_9/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /cgnl\_9/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	64.7	15	21 AAY79270	PHLIX peptide use
2	6	35.3	112	20 AAY30183	Amino acid sequenc
3	6	35.3	112	20 AAY30184	Amino acid sequenc
4	5	29.4	6	4 AAP30295	Sequence of AAs 26
5	5	29.4	6	15 AAR59948	Peptide signal seq
6	5	29.4	9	15 AARG1019	Dynorphin-like pol
7	5	29.4	14	20 AAY42191	Oestrogen receptor
8	5	29.4	16	14 AAR30448	C242:11 Mab kappa
9	5	29.4	16	20 AAY14404	Peptide CDR-L1 der
10	5	29.4	16	21 AAY95222	Anti-platelet glyco
11	5	29.4	21	15 AARG2128	U1 snRNP 70K prote

12	5	29.4	21	16 AAR65087	Random biotinylati
13	5	29.4	21	19 AAW46690	Biotinylation pep
14	5	29.4	21	20 AAY24544	Biotinylation pep
15	5	29.4	25	20 AAY36511	Fragment of human
16	5	29.4	26	20 AAY36630	Fragment of human
17	5	29.4	28	21 AAB37360	Human secreted pep
18	5	29.4	29	20 AAY27910	Human secreted pro
19	5	29.4	30	21 AAB12959	WD domain #4 of hu
20	5	29.4	31	18 AAW24750	Finger 2 domain of
21	5	29.4	33	20 AAY04313	Human secreted pro
22	5	29.4	35	21 AAB02774	Mouse Op-3 finger
23	5	29.4	35	21 AAY92545	Finger 2 subdomain
24	5	29.4	37	21 AAB09507	Murine Op-3 finger
25	5	29.4	38	15 AAR58407	TSAR GAM.9-3 pepti
26	5	29.4	38	21 AAB20677	Type I PKS derived
27	5	29.4	40	16 AAR85049	Peptide rIV from t
28	5	29.4	41	19 AAY86147	S. pneumoniae deri
29	5	29.4	44	21 AAG24877	Arabidopsis thalia
30	5	29.4	46	21 AAG57592	Arabidopsis thalia
31	5	29.4	47	20 AAY11851	Human 5' EST secre
32	5	29.4	49	21 AAG59174	Arabidopsis thalia
33	5	29.4	49	21 AAG61642	Arabidopsis thalia
34	5	29.4	50	20 AAR89341	Streptococcus pneu
35	5	29.4	50	21 AAG55623	Arabidopsis thalia
36	5	29.4	51	21 AAG57591	Arabidopsis thalia
37	5	29.4	52	18 AAW20212	H. pylori transmem
38	5	29.4	52	18 AAW24608	H. pylori transmem
39	5	29.4	53	12 AAR13062	St pre-pro sequenc
40	5	29.4	55	21 AAG02598	Human secreted pro
41	5	29.4	56	21 AAG17385	Arabidopsis thalia
42	5	29.4	56	21 AAG48055	Arabidopsis thalia
43	5	29.4	57	21 AAG03393	Human secreted pro
44	5	29.4	58	21 AAB34806	Human secreted pro
45	5	29.4	59	20 AAY35931	Extended human sec

ALIGNMENTS

RESULT 1  
AAY79270  
ID AAY79270 standard; Peptide; 15 AA.  
XX AC AAY79270;  
DT 03-JUL-2000 (first entry)  
XX DE PHELIX peptide used to raise antibody.  
XX PHELIX; human; testis-specific; transcription factor;  
KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
KW therapy; diagnosis; vaccine; antibody.  
XX OS Homo sapiens.  
XX PN WO200012709-A2.  
XX PD 09-MAR-2000.  
XX PF 31-AUG-1999; 99WO-US20137.  
XX PR 31-AUG-1998; 98US-0098610.  
XX PR 31-OCT-1998; 98US-0106524.  
XX (UROC-) UROGENESYS INC.  
XX PA (AFAR/) AFAR D E.  
XX PA (HUBE/) HUBERT R S.  
XX (RAIT/) RAITANO A B.  
XX PI Afar DE, Hubert RS, Raitano AB;  
XX WPI; 2000-237872/20.

PT Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -

XX Example 5; Page 36; 62pp; English.

XX The present sequence is that of a peptide derived from human

CC PHEIX (see AAY79269), a novel transcription factor that is  
 CC normally expressed only in testis tissue, but which is up-regulated  
 CC in prostate and some other cancers. The peptide was conjugated to  
 CC keyhole limpet haemocyanin and used to raise polyclonal antiserum  
 CC in rabbit. The antiserum demonstrated specificity for PHEIX and  
 CC may therefore be useful for assessing the expression of PHEIX in  
 CC patient samples.

XX Sequence 15 AA;

Query Match 64.7%; Score 11; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.00026;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 HSKKELRRER 17  
 |||||  
 DB 1 hskkellrrer 11

RESULT 2

AAAY30183  
 ID AAY30183 standard; Protein; 112 AA.

XX AC AAY30183;

29-OCT-1999 (first entry)

DE Amino acid sequence of light chain variable region of LO-CD2a.

XX Antibody LO-CD2a; CD2 antigen; T-lymphocyte; humanised antibody;

KW T-cell-mediated immune response; graft rejection; autoimmune disease;  
 KW graft-versus-host disease; T cell; natural killer cell.

XX Rattus sp.

XX US5951983-A.

XX 14-SEP-1999.

XX 07-JUN-1995; 95US-0477989.

XX 07-JUN-1995; 95US-0477989.

XX 05-MAR-1993; 93US-0027008.

XX 09-SEP-1993; 93US-0119032.

XX 29-MAR-1995; 95US-0407009.

XX (BIOT-) BIO TRANSPLANT INC.

PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.

XX Bazin H, Kaplan R, Kieber-Emmons T, Latinne D, Postema CE;

PI White-Scharf ME;

XX WPI; 1999-526991/44.

XX Antibody mediated Inhibition of T cell immune response

XX Disclosure; Column 87; 104pp; English.

XX The present sequence represents the light chain variable region of rat  
 CC monoclonal antibody LO-CD2a. LO-CD2a binds to an epitope of a CD2  
 CC antigen T-lymphocytes. The specification describes humanised LO-CD2a  
 CC antibodies. The humanized LO-CD2a antibody comprises the human constant  
 CC regions, a light chain framework region derived from a human antibody,  
 CC a heavy chain framework region derived from a human antibody, heavy and  
 CC light chain complementarity determining regions (CDRs) of the non-human

CC monoclonal antibody produced by the cell line deposited as ATCC HB11423.  
 CC The humanised antibodies are used in a method for treating a patient to  
 CC inhibit a T-cell-mediated immune response. The method is useful for the  
 CC treatment or prevention of graft rejection and graft-versus-host disease,  
 CC as well as in the treatment of autoimmune diseases which are mediated  
 CC by the activation and proliferation of T cells or natural killer cells.

XX Sequence 112 AA;

Query Match 35.3%; Score 6; DB 20; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHSS 9  
 |||||  
 DB 28 silhss 33

RESULT 3

AAAY30184  
 ID AAY30184 standard; Protein; 112 AA.

XX AC AAY30184;

29-OCT-1999 (first entry)

DE Amino acid sequence of humanised light chain variable region of LO-CD2a.

XX Antibody LO-CD2a; CD2 antigen; T-lymphocyte; humanised antibody;

KW T-cell-mediated immune response; graft rejection; autoimmune disease;  
 KW graft-versus-host disease; T cell; natural killer cell.

XX Synthetic.

OS Rattus sp.

OS Homo sapiens.

XX US5951983-A.

XX 14-SEP-1999.

XX 07-JUN-1995; 95US-0477989.

XX 07-JUN-1995; 95US-0477989.

XX 05-MAR-1993; 93US-0027008.

XX 09-SEP-1993; 93US-0119032.

XX 29-MAR-1995; 95US-0407009.

XX (BIOT-) BIO TRANSPLANT INC.

PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.

XX Bazin H, Kaplan R, Kieber-Emmons T, Latinne D, Postema CE;

PI White-Scharf ME;

XX WPI; 1999-526991/44.

XX Antibody mediated Inhibition of T cell immune response

XX Claim 2; Column 87-89; 104pp; English.

XX The present sequence represents the light chain variable region of a  
 CC humanised rat monoclonal antibody LO-CD2a. LO-CD2a binds to an epitope of  
 CC a CD2 antigen T-lymphocytes. The specification describes humanised  
 CC LO-CD2a antibodies. The humanized LO-CD2a antibody comprises the human  
 CC constant regions, a light chain framework region derived from a human  
 CC antibody, a heavy chain framework region derived from a human antibody,  
 CC heavy and light chain complementarity determining regions (CDRs) of the  
 CC non-human monoclonal antibody produced by the cell line deposited as  
 CC ATCC HB11423. The humanised antibodies are used in a method for treating  
 CC a patient to inhibit a T-cell-mediated immune response. The method is  
 CC useful for the treatment or prevention of graft rejection and  
 CC graft-versus-host disease, as well as in the treatment of autoimmune  
 CC diseases which are mediated by the activation and proliferation of

CC T cells or natural killer cells.

XX Sequence 112 AA;

Query Match 35.3%; Score 6; DB 20; Length 112;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHSS 9

|||||

Db 28 slhss 33

RESULT 4

AAP30295

ID AAP30295 standard; Protein; 6 AA.

XX

AC AAP30295;

XX

DT 20-APR-1992 (first entry)

XX

DE Sequence of AAs 26-31 of the E. coli heat stable toxin

XX which determine an H-epitope.

XX

KW Synthetic vaccine; antigen; allergen; immunological response;

XX antibody.

XX

OS Escherichia coli.

XX

PN EP93851-A.

XX

PD 16-NOV-1983.

XX

PF 11-MAR-1983; 83EP-0102392.

XX

PR 15-MAR-1982; 82US-0358150.

XX

PR 28-JAN-1983; 83US-0461802.

XX

PR 12-JUN-1981; 81US-0272855.

XX

PR 09-JAN-1981; 81US-0223558.

XX

PR 16-DEC-1986; 86US-0942562.

XX

PA (NYBL-) NEW YORK BLOOD CENT.

XX

PI Hopp TP;

XX

DR WPI; 1983-822049/47.

XX

PT Synthetic vaccine - contains peptide residue coupled to higher

XX alkyl or alkenyl Gps. and with 6 amino acids in residue

XX

PS Claim 50; Page 46; 54pp; English.

XX

CC The inventors claim a synthetic vaccine which comprises a peptide  
CC residue coupled to an alkyl or alkenyl gp. having at least 12C, or  
CC other lipophilic substance. The residue contains a sequence of 6 AAs  
CC corresp. to the SQ of such AAs in a protein antigen or allergen where  
CC the greatest local average hydrophilicity is found. Pref. the AAs in  
CC the peptide do not exceed 50 residues, and they esp. contain 12-18  
CC residues. The alkyl or alkenyl gp. pref. contains 12-24C and it is  
CC pref. coupled to the terminal amino gp. of the residue opt. via a CO  
CC gp. If a lipophilic substance is used, it is pref. palmitic, stearic,  
CC behenic, oleic or mycolic acid.

XX Sequence 6 AA;

Query Match

Best Local Similarity 29.4%; Score 5; DB 4; Length 6;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKEX 12

|||||

Db 2 sskex 6

RESULT 5

AAR59948

ID AAR59948 standard; peptide; 6 AA.

XX

AC AAR59948;

XX

DT 14-FEB-1995 (first entry)

XX

DE Peptide signal sequence for treating E. coli infections.

XX

KW Therapeutic; metabolic interactions; PSS; analogues.

XX

OS Synthetic.

XX

PN WO9416328-A.

XX

PD 21-JUL-1994.

XX

PF 30-DEC-1993; 93WO-US12679.

XX

PR 30-DEC-1992; 92US-0997727.

XX

PA (RATH/) RATH M.

XX

PI Rath M;

XX

DR WPI; 1994-249399/30.

XX

PT Identifying peptide signal sequences in a protein - and use of  
PT their synthetic analogues for treating or preventing, e.g.  
PT cardiovascular and auto-immune disease, infections and cancer.

XX

PS Claim 18; Page 10; 28pp; English.

XX

CC The sequence is that of a peptide signal sequence which can be used  
CC to treat E. coli infections.

XX See also AAR59944-83.

XX

SQ Sequence 6 AA;

Query Match

Best Local Similarity 29.4%; Score 5; DB 15; Length 6;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKEX 12

|||||

Db 2 sskex 6

RESULT 6

AAR61019

ID AAR61019 standard; peptide; 9 AA.

XX

AC AAR61019;

XX

DT 23-APR-1995 (first entry)

XX

DE Dynorphin-like polypeptide.

XX

KW Dynorphin; opioid analgesic; stable; stability; intravenous.

XX

OS Synthetic.

XX

PH Key Location/Qualifiers

FT Modified-site 1 /note= "MeTyr"

FT Misc-difference 8 /note= "D-Glu"

FT Modified-site 9

FT XX /note= "Arg-NH2"  
PN EP614913-A.  
XX 14-SEP-1994.  
PD 08-NOV-1985; 94EP-0107769.  
XX PF  
XX 09-NOV-1984; 84JP-0236076.  
PR XX  
XX (EISA ) EISAI CO LTD.  
PA Arakawa Y, Araki S, Ikeda M, Kaneko T, Nakazawa T;  
PI Tachibana S, Tsuchiya Y, Yamatsu K, Yoshino H;  
PI WPI; 1994-281157/35.  
XX  
XX New dynorphin polypeptide derivs. for use as analgesics - contg.  
PT a D-amino acid or an N-alkyl deriv. of a D- or L-amino acid for  
PT in vivo stability.  
XX Example 26; Page 27; 47pp; English.  
PS  
XX The patent discloses dynorphin-like polypeptides which are shorter  
CC than dynorphin (7 to 9 amino acids) and which contain either a D-  
CC amino acid and/or an N-alkyl deriv. of an L- or D-amino acid to  
CC provide in-vivo stability upon intravenous administration. The  
CC peptides have a much greater analgesic effect than dynorphin  
CC due to their greater stability in the blood (e.g. IC50 values can be  
CC as low as 0.04 nM when tested by the rabbit vas deferens method; cf.  
CC 17.4 nM for dynorphin).  
CC The present peptide is a specific example of the peptides.  
XX  
XX Sequence 9 AA;  
SQ  
Query Match 29.4%; Score 5; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 13 LRRER 17  
Db 5 lrrer 9  
RESULT 7  
AAAY42191  
ID AAAY42191 standard; peptide; 14 AA.  
XX  
AC AAAY42191;  
XX  
DT 16-DEC-1999 (first entry)  
XX  
DE Oestrogen receptor-beta MBP peptide M-70.  
XX  
XX Oestrogen receptor beta; ER-beta; MBP; binding; detection;  
KW ligand-dependent nuclear receptor; coactivator interaction;  
KW hybrid system; corepressor; ligand binding domain; LBD.  
XX  
OS Synthetic.  
XX  
PN W09950664-A1.  
XX  
PD 07-OCT-1999.  
XX  
PF 01-APR-1999; 99WO-US07168.  
XX  
XX 01-APR-1998; 98US-0053611.  
XX  
XX (GLAX ) GLAXO GROUP LTD.  
PA Northrop JP, Hart CP, Schatz PJ;  
PI  
XX

DR WPI; 1999-591374/50.  
XX  
PT Identifying agonists or antagonists for ligand-induced activation of  
PT predetermined nuclear receptor having ligand binding domain (LBD) -  
XX  
PS Claim 7; Fig 16; 67pp; English.  
XX  
XX The present invention describes the use of a ligand binding domain (LBD)  
CC transcription factor (TRX) fusion protein in methods for identifying  
CC agonists or antagonists for ligand-induced activation of a predetermined  
CC nuclear receptor having a LBD. Also describes are: (1) a positive hybrid  
CC nuclear receptor signal transduction system, comprising an intact  
CC eukaryotic host cell; (2) a nuclear receptor signal transduction system,  
CC comprising a reverse hybrid reporter host cell; (3) a multiplexed format  
CC assay for identifying nuclear receptor ligands; (4) identifying a  
CC candidate pharmaceutical agent from a library of test agents; (5) a  
CC polypeptide comprising a binding amino acid sequence that is  
CC non-naturally occurring in a nuclear protein, and predetermined to bind  
CC to a nuclear receptor, typically by interaction with a binding  
CC interface of a nuclear receptor at which naturally occurring  
CC coactivators and/or corepressors bind; and (6) an agonist ligand of a  
CC nuclear receptor identified using above methods. The new methods are  
CC useful for identifying agonists or antagonists for ligand-induced  
CC activation of a predetermined nuclear receptor having a LBD. AAAY42186  
CC to AAAY42204 represent specifically claimed examples of (5).  
XX  
SQ Sequence 14 AA;  
Query Match 29.4%; Score 5; DB 20; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 SLLHS 8  
Db 6 silhs 10  
RESULT 8  
AAR30448  
ID AAR30448 standard; peptide; 16 AA.  
XX  
AC AAR30448;  
XX  
DT 06-MAY-1993 (first entry)  
XX  
DE C242:11 MAb kappa chain CDR1.  
XX  
XX Kappa; chain; heavy; complementarity determining region; CDR; MAb;  
KW monoclonal antibody; C242:II; murine; Igg; hybridoma; cell line;  
KW spleen; human; colonic; adenocarcinoma; myeloma; Sp2/O; antigen;  
KW endocytosis.  
XX  
OS Synthetic.  
OS  
PN EP521842-A.  
XX  
XX 07-JAN-1993.  
PD  
XX  
PF 03-JUL-1992; 92EP-0850166.  
XX  
XX 03-JUL-1991; 91SE-0002074;  
PR  
XX (KABI ) KABI PHARMACIA AB.  
PA Holmgren J, Lind P, Lindholm L;  
PI WPI; 1993-002345/01.  
XX  
XX Monoclonal antibody reacting with CA-242 antigen - obt'd. by  
PT culturing hybridoma cell line C242:11 or mutants, useful for  
PT diagnosis and therapy of pancreatic or colorectal cancers  
XX

PS Claim 5; Page 11; 15pp; English.

XX The sequences given in AAR30448-50 represent the kappa chain and  
CC AAR30451-53 the heavy chain complementarity determining regions (CDR)  
CC of a monoclonal antibody (MAB), C242:II. C242:II is a monoclonal  
CC murine Ab of IgG class produced when culturing in an appropriate  
CC medium a hybridoma cell line obtained by fusing spleen cells from a  
CC mouse, which has been immunised with a human colonic adenocarcinoma  
CC cell line, with the murine myeloma cell line Sp2/0. C242:II when  
CC bound to a cell surface antigen is capable of being endocytosed or  
CC internalised into cells.

XX Sequence 16 AA;

Query Match 29.4%; Score 5; DB 14; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLLHS 8  
|||||  
Db 5 silhs 9

RESULT 9

AA114404  
ID AAY14404 standard; peptide; 16 AA.

XX AC AAY14404;

XX 17-AUG-1999 (first entry)

DE Peptide CDR-L1 derived from anti-HCV Ser/Thr protease MAB 8D4.

XX Complementarity determining region; CDR: monoclonal antibody; MAB;  
KW hepatitis C virus; HCV; protease; binding site.

XX Synthetic.

OS JP11127861-A.

XX 18-MAY-1999.

XX 29-OCT-1997; 97JP-0297451.

XX 29-OCT-1997; 97JP-0297451.

XX (NIHA ) JAPAN ENERGY CORP.

XX WPI; 1999-350322/30.

XX Neutralized antibody partial peptide derived from hepatitis C virus  
PT - useful for inhibiting Hepatitis C Virus (HCV) serine protease  
PT activity

PS Example 1; Page 13; 32pp; Japanese.

XX This sequence represents a peptide derived for the sequence of the light  
CC chain variable region complementarity determining region (CDR)-1 of the  
CC anti-hepatitis C virus (HCV) Ser/Thr protease monoclonal antibody (MAB)  
CC 8D4 protein. The invention relates to the use of partial peptides  
CC (AAY14348-Y14353) from the MAB 8D4 for inhibiting HCV serine protease  
CC activity.

XX Sequence 16 AA;

Query Match 29.4%; Score 5; DB 20; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLLHS 8  
|||||

Db 5 silhs 9

RESULT 10

AA195222

ID AAY95222 standard; Peptide; 16 AA.

XX AC AAY95222;

XX 29-AUG-2000 (first entry)

DE Anti-platelet glycoprotein Ib human HIB-5 VL CDR1.

XX Variable light chain; single chain antibody; scFv; human; HIB-5;  
KW glycoprotein Ib alpha; platelet; aggregation; antiaggregant;  
KW antithrombotic; thrombus; therapy; diagnostic; CDR1;  
KW complementarity determining region.

XX Homo sapiens.

XX WO200026667-A1.

XX 11-MAY-2000.

XX 29-OCT-1999; 99WO-US25495.

XX 30-OCT-1998; 98US-0106275.

XX (MILL/) MILLER J L.

XX Miller JL;

XX WPI; 2000-365744/31.

PT Isolated nucleic acid molecule encoding anti-human platelet  
PT glycoprotein Ib alpha molecule useful for producing antibodies which  
PT inhibit platelet aggregation -

PS Claim 19; Fig 8; 89pp; English.

XX The present sequence is that of complementarity determining region  
CC 1 (CDR1) of the light chain variable region (VL) of human single  
CC chain antibody (scFv) HIB-5, which is directed against platelet  
CC glycoprotein Ib (GPIb). The HIB series of scFv antibodies was  
CC isolated from a human synthetic VH and VL scFv library on the  
CC basis of their binding to platelet GPIb. Whether displayed as  
CC surface proteins on a phagemid or secreted as free scFv by  
CC Escherichia coli, the HIB scFv clones are capable of inhibiting  
CC von Willebrand factor-dependent aggregation of platelets. The scFv  
CC are composed of native human protein sequences and are therefore  
CC attractive potential reagents for therapeutic purposes. They  
CC provide a new class of antithrombotic agents, useful for the  
CC prevention of platelet-dependent thrombi in diseased arteries,  
CC bypass grafts, dialysis etc., and can also be used as diagnostic  
CC reagents. Methods of inhibiting aggregation of platelets, of  
CC binding human platelet GPIb alpha and of selecting a VH or VL  
CC region of an antibody that inhibits platelet aggregation are  
CC claimed. Fragments of the scFv VH or VL chain, including CDR  
CC fragments, are also claimed.

XX Sequence 16 AA;

Query Match 29.4%; Score 5; DB 21; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 SLLHS 8

Db 5 silhs 9

RESULT 11

AAR62128  
ID AAR62128 standard; peptide; 21 AA.  
XX  
AC AAR62128;  
XX  
DT 27-APR-1995 (first entry)  
XX  
DE U1 snRNP 70K protein motif similar to HIV-1 gp120/41 protein.  
XX  
KW Small ribonucleoprotein complex; U1 snRNP; 70K protein; epitope;  
KW autoantibody; immunoinfective cluster virus; nuclear protein antigen;  
KW systemic rheumatic disorder; human immunodeficiency virus; HIV-1;  
KW systemic lupus erythematosus; mixed connective tissue disease;  
KW scleroderma.  
XX  
OS Homo sapiens.  
XX  
PN WO9420141-A.  
XX  
PD 15-SEP-1994.  
XX  
PF 10-MAR-1994; 94WO-US02631.  
XX  
PR 11-MAR-1993; 93US-0029850.  
XX  
PA (UYSC-) UNIV SOUTHERN CALIFORNIA.  
XX  
PI Douvas A, Ehresmann G, Takehana Y;  
XX  
DR WPI; 1994-302689/37.  
XX  
PT Methods for treating immunoinfective cluster virus infections -  
PT utilise antibodies or fragments characteristic of auto antibodies  
PT produced by patients with rheumatic disorders  
XX  
PS Disclosure; Page 49; 106pp; English.  
XX  
CC The U1 snRNP is the target of high-titre, high avidity  
CC autoantibodies occurring in the systemic rheumatoid disorders of  
CC mixed connective tissue disease, scleroderma and systemic lupus  
CC erythematosus. It has been found that some sites in the U1 snRNP  
CC 70K protein (see AAR62120-R62135) are homologous to sites in HIV-1  
CC gp120/41 (AAR62136-R62152) and that anti-RNP autoantibodies can be  
CC used to neutralise HIV-1.  
XX  
SQ Sequence 21 AA;

Query Match 29.4%; Score 5; DB 15; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 KLRRE 16  
| | | |  
DB 2 klrre 6

RESULT 12  
AAR65087  
ID AAR65087 standard; Peptide; 21 AA.  
XX  
AC AAR65087;  
XX  
DT 25-OCT-1995 (first entry)  
XX  
DE Random biotinylation peptide 54.  
XX  
KW biotinylation; peptide; recombinant; fusion protein; small;  
KW specific; defined; purification; Bira; enzyme; biotin.  
XX  
OS Synthetic.  
XX  
PI Key

Location/Qualifiers

FT Modified-site 11  
FT /note= "biotin-Lys"  
XX  
PN WO9504069-A.  
XX  
PD 09-FEB-1995.  
XX  
PF 28-JUL-1994; 94WO-US08528.  
XX  
PR 30-JUL-1993; 93US-0099991.  
XX  
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.  
XX  
PI Schatz PJ;  
XX  
DR WPI; 1995-090609/12.  
XX  
PT Production of biotinylated proteins by expression of a  
PT recombinant DNA vector - which encodes a fusion protein  
PT comprising a protein and a biotinylated peptide.  
XX  
PS Claim 10; Page 137; 146pp; English.  
XX  
CC A library of small, efficient peptide biotinylation sequences  
CC (AAR65087-89) was generated by using a generic peptide X10-K-X10 and a  
CC system known as the "peptides on plasmids" system. At some positions in  
CC the sequences, no clear consensus is apparent. At other residues,  
CC however, clear trends emerge. A protein can be biotinylated by  
CC constructing a recombinant DNA expression vector encoding a fusion  
CC protein, comprising a protein and a biotinylation peptide. A host cell,  
CC eg. E. coli is transformed with the vector and is cultured in the  
CC presence of biotin and a biotinylation enzyme, eg. Bira.  
XX  
SQ Sequence 21 AA;

Query Match 29.4%; Score 5; DB 16; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLHLS 8  
| | | |  
DB 5 silhs 9

RESULT 13  
AAW46690  
ID AAW46690 standard; peptide; 21 AA.  
XX  
AC AAW46690;  
XX  
DT 28-MAY-1998 (first entry)  
XX  
DE Biotinylation peptide isolated from random library 4.  
XX  
KW Biotinylation peptide; biotinylation enzyme; biotin-protein ligase;  
KW Bira; biotin ligase; biotin; purification; immobilisation; labelling;  
KW detection; protein.  
XX  
OS Synthetic.  
XX  
PN US5723584-A.  
XX  
PD 03-MAR-1998.  
XX  
PF 03-FEB-1995; 95US-0383753.  
XX  
PR 30-JUL-1993; 93US-0099991.  
XX  
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.  
XX  
PI Schatz PJ;  
XX

DR WPI; 1998-178542/16.  
 XX Peptide(s) that can be biotinylated by biotin ligase - and fusion  
 PT proteins containing them  
 XX  
 XX Claim 3; Column 67; 33pp; English.  
 PS  
 XX Peptides AAW46690-92 are non-naturally occurring biotinylation peptides,  
 CC derived from a library constructed to express peptides of the generic  
 CC sequence:  
 CC (Xaa)10-Lys-(Xaa)10  
 CC where Xaa is a NKK codon, and the codon for Lys is fixed.  
 CC The peptides contain a biotinylatable sequence motif, recognised by a  
 CC biotinylation enzyme, e.g. biotin-protein ligase (BirA). The C or N  
 CC terminus of the peptides can be covalently coupled to a protein that  
 CC is incapable of being biotinylated by a biotin ligase. The peptides  
 CC can be biotinylated in vitro or in vivo, especially with BirA biotin  
 CC ligase, and used for the purification, immobilisation, labelling or  
 CC detection of proteins.  
 XX  
 XX Sequence 21 AA;  
 SQ

Query Match 29.4%; Score 5; DB 19; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHS 8  
 Db 5 silhs 9  
 |||||

RESULT 14  
 AAY24544  
 ID AAY24544 standard; peptide; 21 AA.  
 XX  
 AC AAY24544;  
 XX  
 XX 28-SEP-1999 (first entry)  
 DT  
 DE Biotinylation peptide SEQ ID NO:70.  
 XX  
 KW Biotinylation enzyme; biotin-protein ligase; BirA; labelling;  
 KW biotin carboxyl carrier protein; BCCP; Escherichia coli;  
 KW fusion protein; identification; purification; diagnosis; research.  
 XX  
 OS Escherichia coli.  
 OS Synthetic.  
 XX  
 PN US5932433-A.  
 XX  
 XX 03-AUG-1999.  
 PD  
 XX 28-OCT-1997; 97US-0959512.  
 PF  
 XX 03-FEB-1995; 95US-0383753.  
 PR  
 PR 30-JUL-1993; 93US-0099991.  
 PR  
 PR 28-OCT-1997; 97US-0959512.  
 XX  
 PA (AFY-) AFFYMAX TECHNOLOGIES NV.  
 XX  
 XX Schatz PJ;  
 PI  
 XX WPI; 1999-457113/38.  
 DR  
 XX Identification and biotinylation of proteins synthesized by  
 PT recombinant DNA techniques in vivo  
 PT  
 PS Claim 6; Column 67; 35pp; English.  
 XX  
 XX The present invention describes a method for the identification and  
 CC biotinylation of proteins synthesized by recombinant DNA techniques  
 CC in vivo with a biotinylation peptide of less than 50 amino acids. The

CC method comprises: (a) on the surface of a substrate, providing a fusion  
 CC protein comprising a recombinant protein and a peptide defined by the  
 CC sequence given in AAY24492, where the peptide is capable of being  
 CC biotinylated by a biotin ligase at the lysine residue adjacent to  
 CC position 8 and is 13-50 aa in length; (b) in a predefined region of the  
 CC surface of the substrate, contacting the fusion protein with an enzyme;  
 CC and (c) determining whether the fusion protein has been biotinylated.  
 CC The method allows the identification and purification of biotinylated  
 CC enzymes e.g. BirA. The method is also useful in research and diagnostic  
 CC applications. The method uses small but specific peptides, allowing the  
 CC labelling of a protein at a defined site, this provides improved  
 CC immobilization and avoids the use of antibodies. AAY24493 to AAY24548,  
 CC and AAY29299 to AAY29312 represent specifically claimed examples of  
 CC biotinylation peptides for use in the method of the invention.  
 XX  
 SQ Sequence 21 AA;  
 Query Match 29.4%; Score 5; DB 20; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLHS 8  
 Db 5 silhs 9  
 |||||

RESULT 15  
 AAY36511  
 ID AAY36511 standard; Protein; 25 AA.  
 XX  
 AC AAY36511;  
 XX  
 XX 17-SEP-1999 (first entry)  
 DT  
 DE Fragment of human secreted protein encoded by gene 32.  
 XX  
 KW Human; secreted protein; cancer; tumour; developmental abnormality;  
 KW fetal deficiency; blood disorder; immune system disorder; inflammation;  
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;  
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;  
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;  
 KW digestive disorder; endocrine disorder; infection; AIDS.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO931117-A1.  
 XX  
 XX 24-JUN-1999.  
 PD  
 XX 17-DEC-1998; 98WO-US27059.  
 PF  
 XX 19-DEC-1997; 97US-0068369.  
 PR  
 PR 18-DEC-1997; 97US-0068006.  
 PR  
 PR 18-DEC-1997; 97US-0068007.  
 PR  
 PR 18-DEC-1997; 97US-0068008.  
 PR  
 PR 18-DEC-1997; 97US-0068053.  
 PR  
 PR 18-DEC-1997; 97US-0068054.  
 PR  
 PR 18-DEC-1997; 97US-0068057.  
 PR  
 PR 18-DEC-1997; 97US-0068064.  
 PR  
 PR 18-DEC-1997; 97US-0070923.  
 PR  
 PR 19-DEC-1997; 97US-0068169.  
 PR  
 PR 19-DEC-1997; 97US-0068365.  
 PR  
 PR 19-DEC-1997; 97US-0068367.  
 PR  
 PR 19-DEC-1997; 97US-0068368.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Carter KC, Duan RD, Feng P, Ferrie AM, Florence C;  
 PI Florence K, Greene JM, Janat F, Kyaw H, Moore PA;  
 PI Ni J, Rosen CA, Ruben SM, Shi Y, Soppet DR, Wei Y;  
 PI Yu G;  
 XX

DR WPI; 1999-418749/35.  
XX  
PT New isolated human genes encoding secreted polypeptides  
XX  
PS Disclosure; Page 469; 537pp; English.  
XX  
CC AAX97916 to AAX98029 represent 110 isolated human secreted protein  
CC genes. AAY36224 to AAY36727 represent the secreted proteins encoded by  
CC the 110 human genes. The genes and their corresponding secreted  
CC polypeptides are useful for preventing, treating or ameliorating medical  
CC conditions, e.g. by protein or gene therapy. Also pathological conditions  
CC can be diagnosed by determining the amount of the new polypeptides in a  
CC sample or by determining the presence of mutations in the new genes.  
CC Specific uses are described for each of the 110 genes, based on which  
CC tissues they are most highly expressed in, and include developing  
CC products for the diagnosis or treatment of cancer, tumours, developmental  
CC abnormalities and foetal deficiencies, blood disorders, diseases of the  
CC immune system, autoimmune diseases, inflammation, allergies, Alzheimer's  
CC and cognitive disorders, schizophrenia, arthritis, asthma, psoriasis,  
CC sepsis, skin disorders, atherosclerosis, diabetes, cardiovascular  
CC disorders, kidney disorders, digestive/endocrine disorders, infections  
CC and AIDS. The polypeptides are also useful for identifying their binding  
CC partners. The sequences given in AAX97907 to AAX97915 and AAY36223 are  
CC used in the exemplification of the present invention.  
XX  
SQ Sequence 25 AA;

Query Match 29.4%; Score 5; DB 20; Length 25;  
Best Local Similarity 100.0%; Pred. NO. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KISLL 6  
Db 13 KISLL 17

Search completed: September 15, 2001, 12:51:07  
Job time: 227 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:59:47 ; Search time 8.51 Seconds  
(without alignments)  
68.430 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_150  
Perfect score: 17  
Sequence: 1 KKISLLHSSKEKLRRER 17

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues  
Word size : 0

Total number of hits satisfying chosen parameters: 769

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3	17.6	5	1	BI0B_CITFR
2	3	17.6	10	1	TEMK_RANTE
3	3	17.6	12	1	PA21_MICFM
4	3	17.6	12	1	PA2B_VIPHO
5	3	17.6	12	1	RS16_GINBI
6	3	17.6	12	1	RS19_CLYEP
7	3	17.6	12	1	TM2A_METMA
8	3	17.6	13	1	TEMK_RANTE
9	3	17.6	13	1	TEMK_RANTE
10	3	17.6	13	1	YPE2_LACLC
11	3	17.6	14	1	CAT2_FASHE
12	3	17.6	14	1	MAST_POLJA
13	3	17.6	14	1	MAST_VESCR
14	3	17.6	14	1	MAST_VESLE
15	3	17.6	14	1	PSAG_CUCSA
16	3	17.6	14	1	RS19_PRUPA
17	3	17.6	15	1	CDN5_LITCE
18	3	17.6	15	1	BF1A_MICCR
19	3	17.6	15	1	UN04_PINPS
20	3	17.6	16	1	BP4_PIG
21	2	11.8	4	1	FAR3_HIRME
22	2	11.8	4	1	FLRF_HIRME
23	2	11.8	6	1	ACPH_RABIT
24	2	11.8	6	1	LOK1_LOOMI
25	2	11.8	6	1	THOF_SARBU
26	2	11.8	7	1	CARP_MYTED
27	2	11.8	7	1	CHOX_ALCSP
28	2	11.8	7	1	FAR1_HELTI
29	2	11.8	7	1	FAR1_PROCL
30	2	11.8	7	1	FAR2_ASCSU
31	2	11.8	7	1	FAR2_PROCL
32	2	11.8	7	1	GFRP_MOUSE
33	2	11.8	7	1	HY7_PIG

34	2	11.8	7	1	MSCL_SALTY
35	2	11.8	7	1	MYOM_APLCA
36	2	11.8	7	1	UH11_RAT
37	2	11.8	7	1	WWAL_ACHFU
38	2	11.8	7	1	WWAL_ACHFU
39	2	11.8	8	1	CAD1_ENTFA
40	2	11.8	8	1	FAR1_PANRE
41	2	11.8	8	1	FAR3_HOMAM
42	2	11.8	8	1	FAR4_HOMAM
43	2	11.8	8	1	GLUR_HUMAN
44	2	11.8	8	1	LCK2_LEUMA
45	2	11.8	8	1	LCK4_LEUMA

ALIGNMENTS

RESULT 1					
BI0B_CITFR					
ID BI0B_CITFR	STANDARD;	PRT;	5 AA.		
AC P12997;					
DT 01-JAN-1990 (Rel. 13, Created)					
DT 01-JAN-1990 (Rel. 13, Last sequence update)					
DT 01-OCT-2000 (Rel. 40, Last annotation update)					
DE BIOTIN SYNTHASE (EC 2.8.1.6) (BIOTIN SYNTHETASE) (FRAGMENT).					
GN BI0B.					
OS Citrobacter freundii.					
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;					
OC Citrobacter					
OX NCBI_TaxID=546;					
[1]					
RP SEQUENCE FROM N.A.					
RX MEDLINE=89006280; PubMed=2971595;					
RA Shuan D., Campbell A.					
RT "Transcriptional regulation and gene arrangement of Escherichia coli,					
RT Citrobacter freundii and Salmonella typhimurium biotin operons.";					
RL Gene 67:203-211(1988).					
CC - - CATALYTIC ACTIVITY: DETHIOBIOTIN + (S) - BIOTIN.					
CC - - PATHWAY: LAST STEP IN BIOTIN BIOSYNTHESIS PATHWAY.					
CC - - SIMILARITY: BELONGS TO THE BIOTIN AND LIPOIC ACID SYNTHETASES FAMILY.					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC the European Bioinformatics Institute. There are no restrictions on its					
CC use by non-profit institutions as long as its content is in no way					
CC modified and this statement is not removed. Usage by and for commercial					
CC entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).					
CC -----					
DR EMBL; M21922; ; NOT ANNOTATED_CDS.					
KW Biotin biosynthesis; Iron-sulfur; Transferase.					
FT NON_TER 5					
SQ SEQUENCE 5 AA; 532 MW; 75A5B1EDD6F00000 CRC64;					

Query Match 17.6%; Score 3; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	7	HSS	9
	111		
Db	3	HSS	5

RESULT 2					
TEMK_RANTE					
ID TEMK_RANTE	STANDARD;	PRT;	10 AA.		
AC P56923;					
DT 30-MAY-2000 (Rel. 39, Created)					
DT 01-OCT-2000 (Rel. 40, Last sequence update)					
DT 01-OCT-2000 (Rel. 40, Last annotation update)					
DE TEMPORIN K.					

OS Rana temporaria (European common frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.  
 OX NCBI\_TaxID=8407;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Skin;  
 RX MEDLINE=97175050; PubMed=9022710;  
 RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,  
 RA Barra D.;  
 RT "temporins, antimicrobial peptides from the European red frog Rana  
 temporaria.";  
 RL Eur. J. Biochem. 242:788-792(1996).  
 CC -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-POSITIVE  
 CC BACTERIA.  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GARGURIN/RUGOSIN  
 CC FAMILY.  
 KW Amphibian skin; Antibiotic; Amidation; Multigene family.  
 FT MOD\_RES 10 10 AMIDATION.  
 SQ SEQUENCE 10 AA; 1123 MW; 390549B337272457 CRC64;

Query Match 17.6%; Score 3; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLL 6  
 |||  
 DB 8 SLL 10

RESULT 3  
 ID PA21\_MICFM STANDARD; PRT; 12 AA.  
 AC P25072;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE PHOSPHOLIPASE A2 ISOZYME 1 (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE  
 DE 2-ACYLHYDROLASE) (FRAGMENT).  
 OS Micrurus fulvius microgalbineus (Mexican coral snake).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;  
 OC Elapidae; Elapinae; Micrurus.  
 OX NCBI\_TaxID=8636;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Venom;  
 RX MEDLINE=7925521; PubMed=475771;  
 RA Posani L.D., Alagon A.C., Fletcher P.L. Jr., Varela M.J., Julia J.Z.;  
 RT "Purification and characterization of a phospholipase A2 from the  
 RT venom of the coral snake, Micrurus fulvius microgalbineus (Brown and  
 RT Smith).";  
 RL Biochem. J. 179:603-606(1979).  
 CC -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE  
 CC 2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.  
 CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 1-ACYLGLYCERYL-  
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.  
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.  
 DR InterPro; IPR001211;  
 DR PROSITE; PS00118; PA2\_HIS; PARTIAL.  
 DR PROSITE; PS00119; PA2\_ASP; PARTIAL.  
 KW Hydrolase; Lipid degradation; Calcium; Venom.  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1398 MW; CC21992A899F0339 CRC64;

Query Match 17.6%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLL 6

Db 1 SLL 3  
 |||  
 RESULT 4  
 ID PA2B\_VIPBO STANDARD; PRT; 12 AA.  
 AC P31859;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE PHOSPHOLIPASE A2, BASIC (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE  
 DE 2-ACYLHYDROLASE) (FRAGMENT).  
 OS Vipera berus orietalis (Viper).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;  
 OC Viperidae; Viperinae; Vipera.  
 OX NCBI\_TaxID=31157;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Venom;  
 RX MEDLINE=74128698; PubMed=4206446;  
 RA Delori P.J.;  
 RT "Purification and physicochemical, chemical and biological properties  
 RT of a toxic A2 phospholipase isolated from the venom of viperidae  
 RT snakes: Vipera berus.";  
 RL Biochimie 55:1031-1045(1973).  
 CC -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE  
 CC 2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.  
 CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 1-ACYLGLYCERYL-  
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.  
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.  
 DR InterPro; IPR001211;  
 DR PROSITE; PS00118; PA2\_HIS; PARTIAL.  
 DR PROSITE; PS00119; PA2\_ASP; PARTIAL.  
 KW Hydrolase; Lipid degradation; Calcium; Venom.  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1412 MW; D054351A89969879 CRC64;

Query Match 17.6%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLL 6  
 |||  
 DB 1 SLL 3

RESULT 5  
 ID RR16\_GINBI STANDARD; PRT; 12 AA.  
 AC P36207;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S16 (FRAGMENT).  
 GN RPS16.  
 OS Ginkgo biloba (Ginkgo).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Ginkgoophyta; Ginkgoales; Ginkgoaceae; Ginkgo.  
 OX NCBI\_TaxID=3311;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95094313; PubMed=8001171;  
 RA Richard M., Tremblay C., Bellemare G.;  
 RT "Chloroplastic genomes of Ginkgo biloba and Chlamydomonas moewusii  
 RT contain a chlB gene encoding one subunit of a light-independent  
 RT protochlorophyllide reductase.";  
 RL Curr. Genet. 26:159-165(1994).  
 CC -!- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.  
 -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U01531; AAA66977.1; -

DR InterPro; IPR000307; -

DR PROSITE; PS00732; RIBOSOMAL\_S16; 1.

KW Ribosomal protein; Chloroplast.

FT NON\_TER 12

SQ SEQUENCE 12 AA; 1488 MW; 6700EDAF9D033734 CRC64;

Query Match 17.6%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 KLR 14

DB 3 KLR 5

RESULT 6

ID RS19\_CLYPEP

AC Q46490; STANDARD; PRT; 12 AA.

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DE 30S RIBOSOMAL PROTEIN S19 (FRAGMENT).

GN RPS5 OR RPS19.

OS Clover yellow edge phytoplasma.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;

OC Achleplasmataceae; Phytoplasma.

OX NCBI\_TaxID=35775;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94350802; PubMed=80711198;

RA Gundersen D.E., Lee I.M., Rehner S.A., Davis R.E., Kingsbury D.T.;

RT "Phylogeny of mycoplasma-like organisms (phytoplasmas): a basis for

their classification.";

RL J. Bacteriol. 176:5244-5254(1994).

CC -!- FUNCTION: PROTEIN S19 FORMS A COMPLEX WITH S13 THAT BINDS STRONGLY

TO THE 16S RIBOSOMAL RNA (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; L27019; AAA83940.1; -

DR InterPro; IPR002222; -

DR PROSITE; PS00323; RIBOSOMAL\_S19; PARTIAL.

KW Ribosomal protein; rRNA-binding.

FT NON\_TER 1

SQ SEQUENCE 12 AA; 1409 MW; 4CC478EFC8333AA3 CRC64;

Query Match 17.6%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKI 3

DB 7 KKI 9

RESULT 7  
 TM2A\_METMA  
 ID TM2A\_METMA STANDARD; PRT; 12 AA.  
 AC P80652;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE ALTERNATIVE TETRAHYDROMETHANOPTERIN S-METHYLTRANSFERASE 28 KDA SUBUNIT  
 (EC 2.1.1.86) (N5-METHYLTETRAHYDROMETHANOPTERIN-COENZYME M  
 METHYLTRANSFERASE 28 KDA SUBUNIT) (FRAGMENT).  
 OS Methanosarcina mazel.  
 OC Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;  
 OC Methanosarcina.  
 OX NCBI\_TaxID=2209;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=DSM 3647 / GOEL.  
 RX MEDLINE=96370840; PubMed=8774736;  
 RA Lienard T., Becher B., Marschall M., Bowlen S., Gottschalk G.;  
 RT "Sodium ion translocation by N5-methyltetrahydromethanopterin:  
 coenzyme M methyltransferase from Methanosarcina mazel Gol  
 reconstituted in ether lipid liposomes.";  
 RL Eur. J. Biochem. 239:857-864(1996).  
 CC -!- FUNCTION: THIS ENZYME COMPLEX CATALYZES AN INTERMEDIATE STEP IN  
 METHANOGENESIS, THE FORMATION OF METHYL-COENZYME M AND  
 TETRAHYDROMETHANOPTERIN FROM COENZYME M AND N5-METHYL-  
 TETRAHYDROMETHANOPTERIN.  
 CC -!- CATALYTIC ACTIVITY: 5-METHYL-5,6,7,8-TETRAHYDROMETHANOPTERIN +  
 2-MERCAPTOETHANESULFONATE -> 5,6,7,8-TETRAHYDROMETHANOPTERIN +  
 2-(METHYLTHIO)ETHANESULFONATE.  
 CC -!- SUBUNIT: COMPOSED OF SIX DIFFERENT SUBUNITS.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 KW Transferase; Methyltransferase; Transmembrane; Methanogenesis.  
 FT NON\_TER 12  
 SQ SEQUENCE 12 AA; 1321 MW; 6DE4A5766232D76B CRC64;

Query Match 17.6%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 EKL 13

DB 2 EKL 4

RESULT 8

TEM2D\_RANTE

ID TEM2D\_RANTE STANDARD; PRT; 13 AA.

AC P56919;

DT 30-MAY-2000 (Rel. 39, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE TEMPORIN D.

OS Rana temporaria (European common frog).

OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Rana.

OX NCBI\_TaxID=8407;

RN [1]

RP SEQUENCE, AND SYNTHESIS.

RC TISSUE=Skin;

RX MEDLINE=97175050; PubMed=9022710;

RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,

RA Barra D.;

RT "Temporins, antimicrobial peptides from the European red frog Rana

temporaria.";

RL Eur. J. Biochem. 242:788-792(1996).

CC -!- FUNCTION: HAS NO ANTIBACTERIAL ACTIVITY.

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN

FAMILY.

CC Amphibian skin; Amidation; Multigene family.

```

FT  MOD_RES      13      13      AMIDATION.
SQ  SEQUENCE     13 AA; 1379 MW; 3EF35DFA655B2448 CRC64;

Query Match
Best Local Similarity 17.6%; Score 3; DB 1; Length 13;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  4 SLL 6
DB  11 SLL 13

RESULT 9
TEME_RANTE      STANDARD;      PRT;      13 AA.
AC  P56920;
DT  30-MAY-2000 (Rel. 39, Created)
DT  01-OCT-2000 (Rel. 40, Last sequence update)
DT  01-OCT-2000 (Rel. 40, Last annotation update)
DE  TEMPORIN E.
OS  Eura temporaria (European common frog).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX  NCBI_TaxID=8407;
RN  [1]
RP  SEQUENCE.
RC  TISSUE=Skin;
RX  MEDLINE=97175050; PubMed=9022710;
RA  Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA  Barra D.;
RT  "Temporins, antimicrobial peptides from the European red frog Rana
RT  temporaria."
RL  Eur. J. Biochem. 242:788-792(1996).
CC  -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-POSITIVE
CC  BACTERIA.
CC  -!- SUBCELLULAR LOCATION: SECRETED.
CC  -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC  FAMILY.
KW  Amphibian skin; Antibiotic; Amidation; Multigene family.
FT  MOD_RES      13      13
SQ  SEQUENCE     13 AA; 1379 MW; 26505DFA79A92448 CRC64;

Query Match
Best Local Similarity 17.6%; Score 3; DB 1; Length 13;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  4 SLL 6
DB  11 SLL 13

RESULT 10
YPE2_LACLC
ID  YPE2_LACLC      STANDARD;      PRT;      13 AA.
AC  P42021;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  01-NOV-1995 (Rel. 32, Last annotation update)
DE  HYPOTHETICAL PROTEIN IN PEPT 5'REGION (ORF2) (FRAGMENT).
OS  Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OC  Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC  Lactococcus.
OX  NCBI_TaxID=1359;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=94245610; PubMed=8188586;
RA  Mierau I., Haandrikman A.J., Velterop O., Tan P.S.T., Leenhouts K.L.,
RA  Konings W.N., Venema G., Kok J.;
RT  "Tripeptidase gene (pept) of Lactococcus lactis: molecular cloning
RT  and nucleotide sequencing of pept and construction of a chromosomal
RT  deletion mutant."

```

---

```

RL  J. Bacteriol. 176:2854-2861(1994).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; L27596; AAA20625.1; -
KW  Hypothetical protein.
FT  NON_TER      1
SQ  SEQUENCE     13 AA; 1392 MW; 3671D53271B459D7 CRC64;

Query Match
Best Local Similarity 17.6%; Score 3; DB 1; Length 13;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  3 ISL 5
DB  8 ISL 10

RESULT 11
CAT2_FASHE
ID  CAT2_FASHE      STANDARD;      PRT;      14 AA.
AC  P80342;
DT  01-OCT-1994 (Rel. 30, Created)
DT  01-OCT-1994 (Rel. 30, Last sequence update)
DT  01-NOV-1995 (Rel. 32, Last annotation update)
DE  CATHEPSIN L2 (EC 3.4.22.15) (FRAGMENT).
OS  Fasciola hepatica (Liver fluke).
OC  Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
OC  Trematoda; Digenea; Echinostomida; Echinostomata; Fascioloidae;
OC  Fasciolidae; Fasciola.
OX  NCBI_TaxID=6192;
RN  [1]
RP  SEQUENCE.
RX  MEDLINE=94307282; PubMed=8033913;
RA  Dowd A.J., Smith A.M., McGonicle S., Dalton J.P.;
RT  "Purification and characterisation of a second cathepsin L proteinase
RT  secreted by the parasitic trematode Fasciola hepatica.";
RL  Eur. J. Biochem. 223:91-98(1994).
CC  -!- FUNCTION: THIOL PROTEASE THAT ASSISTS THE PARASITE IN BURROWING
CC  THROUGH THE GUT WALL AND LIVER OF ITS MAMMALIAN HOST.
CC  -!- SUBUNIT: DIMER OF AN HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE
CC  BONDS.
CC  -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC  -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
CC  PAPAIN FAMILY OF THIOL PROTEASES.
DR  MEROPS; C01.033; -
DR  InterPro; IPR000169; -
DR  PROSITE; PS00139; THIOL PROTEASE CYS; PARTIAL.
DR  PROSITE; PS00639; THIOL PROTEASE HIS; PARTIAL.
DR  PROSITE; PS00640; THIOL PROTEASE ASN; PARTIAL.
KW  Hydrolase; Thiol protease; Lysosome.
FT  NON_TER      14
SQ  SEQUENCE     14 AA; 1605 MW; 9CAEAB74E9DA110A CRC64;

Query Match
Best Local Similarity 17.6%; Score 3; DB 1; Length 14;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  14 RRE 16
DB  8 RRE 10

RESULT 12
MAST_POLJA

```

```

ID  MAST_POLJA  STANDARD;  PRT;  14 AA.
AC  P01517;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  01-AUG-1992 (Rel. 23, Last annotation update)
DE  POLISTES MASTOPARAN
OS  Polistes jadwigae (Paper wasp).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC  Vespoidea; Vespidae; Polistinae; Polistes.
OX  NCBI_TaxID=7457;
RN  [1]
RP  SEQUENCE.
RC  TISSUE=Venom;
RA  Hirai Y., Ueno Y., Yasuhara T., Yoshida H., Nakajima T.;
RT  "A new mast cell degranulating peptide, polistes mastoparan, in the
RL  venom of Polistes jadwigae.";
RL  Biomed. Res. 1:185-187(1980).
CC  -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. ACTIVATES G PROTEINS
CC  THAT COUPLE TO PHOSPHOLIPASE C.
DR  PIR: A01780; QMWAPP.
KW  Mast cell degranulation; Venom; Amidation.
FT  MOD_RES 14 14
SQ  SEQUENCE 14 AA; 1636 MW; 26472A53BF4778D8 CRC64;

Query Match 17.6%; Score 3; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKI 3
Db 4 KKI 6

RESULT 13
MAST_VESCR
ID  MAST_VESCR  STANDARD;  PRT;  14 AA.
AC  P01516;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  01-AUG-1992 (Rel. 23, Last annotation update)
DE  MASTOPARAN C.
OS  Vespa crabro (European hornet).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC  Vespoidea; Vespidae; Vespinae; Vespa.
OX  NCBI_TaxID=7445;
RN  [1]
RP  SEQUENCE.
RC  TISSUE=Venom;
RA  Argiolas A., Pisano J.J.;
RT  "Isolation and characterization of two new peptides, mastoparan C and
RT  crabrolin, from the venom of the European hornet, Vespa crabro.";
RL  J. Biol. Chem. 259:10106-10111(1984).
CC  -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. ACTIVATES G PROTEINS
CC  THAT COUPLE TO PHOSPHOLIPASE C.
DR  PIR: A01779; QMVHP2.
KW  Mast cell degranulation; Venom; Amidation.
FT  MOD_RES 14 14
SQ  SEQUENCE 14 AA; 1508 MW; 550C0ECALD6AB1D7 CRC64;

Query Match 17.6%; Score 3; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKI 3
Db 11 KKI 13

```

```

RESULT 14
MAST_VESLE
ID  MAST_VESLE  STANDARD;  PRT;  14 AA.
AC  P01514;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  01-JUN-1994 (Rel. 29, Last annotation update)
DE  MASTOPARAN.
OS  Vespusula lewisii (Yellow jacket) (Wasp).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC  Vespoidea; Vespidae; Vespinae; Vespula.
OX  NCBI_TaxID=7452;
RN  [1]
RP  SEQUENCE, AND SYNTHESIS.
RC  TISSUE=Venom;
RX  MEDLINE=80155337; PubMed=540362;
RA  Hirai Y., Yasuhara T., Yoshida H., Nakajima T., Fujino M., Kitada C.;
RT  "A new mast cell degranulating peptide 'mastoparan' in the venom of
RT  Vespusula lewisii.";
RL  Chem. Pharm. Bull. 27:1942-1944(1979).
CC  -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. ACTIVATES G PROTEINS
CC  THAT COUPLE TO PHOSPHOLIPASE C.
DR  PIR: A01776; QMWAV.
KW  Mast cell degranulation; Venom; Amidation.
FT  MOD_RES 14 14
SQ  SEQUENCE 14 AA; 1480 MW; CC0C0ECALD7B00DD CRC64;

Query Match 17.6%; Score 3; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKI 3
Db 11 KKI 13

RESULT 15
PSAG_CUCSA
ID  PSAG_CUCSA  STANDARD;  PRT;  14 AA.
AC  P42049;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  PHOTOSYSTEM I REACTION CENTRE SUBUNIT V (PHOTOSYSTEM I 6.8 KDA
DE  PROTEIN) (PSI-G) (PS I SUBUNIT 10) (FRAGMENT).
GN  PSAG.
OS  Cucumis sativus (Cucumber).
OC  Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC  Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC  Cucurbitales; Cucurbitaceae; Cucumis.
OX  NCBI_TaxID=3659;
RN  [1]
RP  SEQUENCE.
RC  TISSUE=Cotyledon;
RX  MEDLINE=91355209; PubMed=1883835;
RA  Iwasaki Y., Ishikawa H., Hibino T., Takabe T.;
RT  "Characterization of genes that encode subunits of cucumber PS I
RT  complex by N-terminal sequencing.";
RL  Biochim. Biophys. Acta 1059:141-148(1991).
CC  -!- FUNCTION: NOT YET KNOWN.
CC  -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. CHLOROPLAST
CC  THYLAKOID MEMBRANE; ASSOCIATED WITH THE STROMAL SIDE OF THE
CC  THYLAKOID MEMBRANE.
DR  InterPro: IPR000549; -.
DR  PROSITE: PS01026; PHOTOSYSTEM_I_PSAKG; PARTIAL.
KW  Photosynthesis; Photosystem I; Chloroplast; Thylakoid membrane.
FT  NON_TER 14 14
SQ  SEQUENCE 14 AA; 1452 MW; D712F0B91F1CA142 CRC64;

Query Match 17.6%; Score 3; DB 1; Length 14;

```

Best Local Similarity 100.0%; Pred..No. 2.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ISL 5  
| | |  
Db 8 ISL 10

Search completed: September 15, 2001, 13:01:23  
Job time: 96 sec.

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 13:00:02 ; Search time 17.57 Seconds  
(without alignments)  
128.013 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_150

Perfect score: 17

Sequence: 1 KKISLLHSSKEKLRER 17

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2841

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_15.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phage.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_unclassified.\*

13: sp\_vertebrate.\*

14: sp\_virus.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	23.5	12	2 Q9R3B3	Q9R3B3 helicobacte
2	4	23.5	13	2 Q47693	Q47693 escherichia
3	4	23.5	14	11 Q9Z0G5	Q9Z0G5 mus musculu
4	4	23.5	15	5 Q9TWP9	Q9TWP9 hirudo medi
5	4	23.5	16	4 Q9UBI5	Q9UBI5 homo sapien
6	4	23.5	16	4 Q9UD21	Q9UD21 homo sapien
7	3	17.6	7	14 Q66205	Q66205 porcine tra
8	3	17.6	8	6 O18854	O18854 canis famil
9	3	17.6	8	8 Q9XNP8	Q9XNP8 boophilus m
10	3	17.6	8	9 Q37854	Q37854 bacterioph
11	3	17.6	9	4 Q95574	Q95574 homo sapien
12	3	17.6	9	4 Q9UE26	Q9UE26 homo sapien
13	3	17.6	9	6 Q9GK05	Q9GK05 capra hircu
14	3	17.6	10	2 Q60192	Q60192 spiroplasma
15	3	17.6	10	2 Q60194	Q60194 spiroplasma
16	3	17.6	10	2 Q51812	Q51812 escherichia
17	3	17.6	10	4 Q60912	Q60912 homo sapien
18	3	17.6	10	8 Q35013	Q35013 meloidogyne
19	3	17.6	11	2 Q47345	Q47345 escherichia

20	3	17.6	11	2 Q9R7U8	Q9R7U8 pseudomonas
21	3	17.6	11	2 Q9K332	Q9K332 staphylococ
22	3	17.6	11	4 Q94785	Q94785 homo sapien
23	3	17.6	11	5 Q9VN99	Q9VN99 drosophila
24	3	17.6	11	8 Q34380	Q34380 drosophila
25	3	17.6	11	8 Q9G5Y3	Q9G5Y3 agama bibro
26	3	17.6	11	10 Q04131	Q04131 lycopersico
27	3	17.6	11	10 Q9S8X4	Q9S8X4 glycine max
28	3	17.6	11	11 Q9QVC5	Q9QVC5 rattus norv
29	3	17.6	12	2 Q50303	Q50303 bacillus st
30	3	17.6	12	2 P82907	P82907 bacillus li
31	3	17.6	12	4 Q9UMR0	Q9UMR0 homo sapien
32	3	17.6	12	4 Q9UMQ9	Q9UMQ9 homo sapien
33	3	17.6	12	6 Q9XT42	Q9XT42 canis famil
34	3	17.6	12	6 Q9TQW3	Q9TQW3 bos taurus
35	3	17.6	12	7 Q07884	Q07884 pan troglod
36	3	17.6	12	11 Q54970	Q54970 mus musculu
37	3	17.6	12	11 Q9QZD0	Q9QZD0 mus musculu
38	3	17.6	12	13 P82820	P82820 rana catesb
39	3	17.6	12	14 O10421	O10421 influenza a
40	3	17.6	13	2 Q55234	Q55234 synechocyst
41	3	17.6	13	2 Q9RG00	Q9RG00 mycoplasma
42	3	17.6	13	2 Q9REJ7	Q9REJ7 actinobacil
43	3	17.6	13	4 Q14462	Q14462 homo sapien
44	3	17.6	13	4 Q75905	Q75905 homo sapien
45	3	17.6	13	4 Q9UM46	Q9UM46 homo sapien

## ALIGNMENTS

RESULT	1				
Q9R3B3					
ID	Q9R3B3	PRELIMINARY;	PRT;	12 AA.	
AC	Q9R3B3;				
DT	01-MAY-2000 (TREMBLrel. 13, Created)				
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)				
DE	UREASE LARGE SUBUNIT (FRAGMENT).				
OS	Helicobacter felis, and Helicobacter pylori (Campylobacter pylori).				
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;				
OC	Helicobacter.				
OX	NCBI_TaxID=214, 210;				
RN	[1]				
RP	SEQUENCE.				
RX	MEDLINE=93084378; PubMed=1452359;				
RA	Turbett G.R., Hoj P.B., Horne R., Mee B.J.;				
RT	"Purification and Characterization of the urease enzymes of				
RT	Helicobacter species from humans and animals.";				
RL	Infect. Immun. 60:5259-5266(1992).				
SQ	SEQUENCE 12 AA; 1500 MW; 93F05D2362132415 CRC64;				

Query Match 23.5%; Score 4; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	KKIS 4
Db	2	KKIS 5

RESULT	2				
Q47693					
ID	Q47693	PRELIMINARY;	PRT;	13 AA.	
AC	Q47693;				
DT	01-NOV-1996 (TREMBLrel. 01, Created)				
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)				
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)				
DE	ELONGATION FACTOR TU (EF-TU) (FRAGMENT).				
GN	TUFB.				
OS	Escherichia coli.				
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				

OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82080657; PubMed=7312036;  
 RA Hudson L., Rossi J., Landy A.;  
 RT "Dual function transcripts specifying tRNA and mRNA.";  
 RL Nature 294:422-427(1981).  
 DR EMBL; X04181; CAA27777.1; -;  
 KW Elongation factor; Protein biosynthesis.  
 FT NON\_TER 13  
 SQ SEQUENCE 13 AA; 1617 MW; C433BE82A18D0B19 CRC64;  
  
 Query Match 23.5%; Score 4; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 9 SKEK 12  
 Db 2 SKEK 5  
 ||||  
  
 RESULT 3  
 Q9Z0G5  
 ID Q9Z0G5 PRELIMINARY; PRT; 14 AA.  
 AC Q9Z0G5;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE SHORTER ISOFORM OF INTERLEUKIN 15 (FRAGMENT).  
 GN INTERLEUKIN 15.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALBC;  
 RA Nishimura H.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALBC;  
 RX MEDLINE=98211658; PubMed=9551932;  
 RA Nishimura H., Washizu J., Nakamura N., Enomoto A., Yoshikai Y.;  
 RT "Translational efficiency is up-regulated by alternative exon in  
 RT murine IL-15 mRNA.";  
 RL J. Immunol. 160:936-942(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Prinz M., Hanisch U.K., Kettenmann H., Kirchhoff P.;  
 RT "Alternative splicing of mouse IL-15 is due to the use of an internal  
 RT splice site in exon 5.";  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB022307; BAA37122.1; -;  
 DR EMBL; AJ012587; CAA10069.1; -;  
 FT NON\_TER 14  
 SQ SEQUENCE 14 AA; 1503 MW; 62A43C3889924E2B CRC64;  
  
 Query Match 23.5%; Score 4; DB 11; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 8 SKEK 11  
 Db 4 SKEK 7  
 ||||  
  
 RESULT 4  
 Q9TWP9  
 ID Q9TWP9 PRELIMINARY; PRT; 15 AA.

AC Q9TWP9;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
 DE G39 ANTIGEN (FRAGMENT).  
 OS Hirudo medicinalis (Medicinal leech).  
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;  
 OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.  
 OX NCBI\_TaxID=6421;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=94157531; PubMed=8113784;  
 RA Luthi T.E., Brodbeck D.L., Jeno P.;  
 RT "Identification of a 70 kD protein with sequence homology to squid  
 RT neurofilament protein in glial cells of the leech CNS.";  
 RL J. Neurobiol. 25:70-82(1994).  
 SQ SEQUENCE 15 AA; 1737 MW; 1E35B4884A11C95E CRC64;  
  
 Query Match 23.5%; Score 4; DB 5; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 3 ISLL 6  
 Db 12 ISLL 15  
 ||||  
  
 RESULT 5  
 Q9UBI5  
 ID Q9UBI5 PRELIMINARY; PRT; 16 AA.  
 AC Q9UBI5;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE CYSTATIN S.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92138674; PubMed=1778989;  
 RA Isemura S., Saitoh E., Sanada K., Minakata K.;  
 RT "Identification of full-sized forms of salivary (S-type) cystatins  
 RT (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of  
 RT cystatin S) in human whole saliva and determination of phosphorylation  
 RT sites of cystatin S.";  
 RL J. Biochem. 110:648-654(1991).  
 RN [2]  
 RP SEQUENCE.  
 RX MEDLINE=92074898; PubMed=1741693;  
 RA Johnsson M., Richardson C.F., Bergey E.J., Levine M.J.,  
 RA Nancollas G.H.;  
 RT "The effects of human salivary cystatins and statherin on  
 RT hydroxyapatite crystallization.";  
 RL Arch. Oral Biol. 36:631-636(1991).  
 SQ SEQUENCE 16 AA; 1765 MW; A6EDF31C3186F9C9 CRC64;  
  
 Query Match 23.5%; Score 4; DB 4; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 8 SKEK 11  
 Db 2 SKEK 5  
 ||||  
  
 RESULT 6  
 Q9UD21  
 ID Q9UD21 PRELIMINARY; PRT; 16 AA.  
 AC Q9UD21;



```
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE CYCLIN E-L (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95257942; PubMed=7739542;
RA Ohtsubo M., Theodoras A.M., Schumacher J., Roberts J.M., Pagano M.;
RT "Human cyclin E, a nuclear protein essential for the G1-to-S phase
transition."
RL Mol. Cell. Biol. 15:2612-2624(1995).
SQ SEQUENCE 16 AA; 2089 MW; 777EFC69C445E29C CRC64;

Query Match 23.5%; Score 4; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 RRER 17
DB 5 RRER 8

RESULT 7
Q66205
ID Q66205 PRELIMINARY; PRT; 7 AA.
AC Q66205;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE MEMBRANE PROTEIN (1 IS 3RD BASE IN CODON) (FRAGMENT).
OS Porcine transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=FS772/70; PubMed=2835592;
RA Britton P., Carmones R.S., Page K.W., Garwes D.J., Parra F.;
RT "Sequence of the nucleoprotein gene from a virulent British field
isolate of transmissible gastroenteritis virus and its expression in
Saccharomyces cerevisiae."
RL Mol. Microbiol. 2:89-99(1988).
DR EMBL; Y00342; CAA68606.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 869 MW; 72C691E727233B70 CRC64;

Query Match 17.6%; Score 3; DB 14; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 EKL 13
DB 1 EKL 3

RESULT 8
O18854
ID O18854 PRELIMINARY; PRT; 8 AA.
AC O18854;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)
DE CD19 ANTIGEN (FRAGMENT).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE CYCLIN E-L (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95257942; PubMed=7739542;
RA Ohtsubo M., Theodoras A.M., Schumacher J., Roberts J.M., Pagano M.;
RT "Human cyclin E, a nuclear protein essential for the G1-to-S phase
transition."
RL Mol. Cell. Biol. 15:2612-2624(1995).
SQ SEQUENCE 16 AA; 2089 MW; 777EFC69C445E29C CRC64;

Query Match 23.5%; Score 4; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 RRER 17
DB 5 RRER 8

RESULT 7
Q66205
ID Q66205 PRELIMINARY; PRT; 7 AA.
AC Q66205;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE MEMBRANE PROTEIN (1 IS 3RD BASE IN CODON) (FRAGMENT).
OS Porcine transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=FS772/70; PubMed=2835592;
RA Britton P., Carmones R.S., Page K.W., Garwes D.J., Parra F.;
RT "Sequence of the nucleoprotein gene from a virulent British field
isolate of transmissible gastroenteritis virus and its expression in
Saccharomyces cerevisiae."
RL Mol. Microbiol. 2:89-99(1988).
DR EMBL; Y00342; CAA68606.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 869 MW; 72C691E727233B70 CRC64;

Query Match 17.6%; Score 3; DB 14; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 EKL 13
DB 1 EKL 3

RESULT 8
O18854
ID O18854 PRELIMINARY; PRT; 8 AA.
AC O18854;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)
DE CD19 ANTIGEN (FRAGMENT).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
```

```
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu P.-C., Shibuya H., Katz M.L., Johnson G.S.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF024717; AAB81967.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 832 MW; 6735A1ADD81325A7 CRC64;

Query Match 17.6%; Score 3; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 SKE 11
DB 2 SKE 4

RESULT 9
Q9XNP8
ID Q9XNP8 PRELIMINARY; PRT; 8 AA.
AC Q9XNP8;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE ATP SYNTHASE 6 (FRAGMENT).
GN ATP6.
OS Boophilus microplus (Cattle tick).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Boophilus.
OX NCBI_TaxID=6941;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N;
RX MEDLINE=99297341; PubMed=10368952;
RA Campbell N.J.H., Barker S.C.;
RT "The novel mitochondrial gene arrangement of the cattle tick,
Boophilus microplus: fivefold tandem repetition of a coding region."
RL Mol. Biol. Evol. 16:732-740(1999).
DR EMBL; AF110616; AAD28386.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 938 MW; 58BB14404B5735B0 CRC64;

Query Match 17.6%; Score 3; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ISL 5
DB 1 ISL 3

RESULT 10
Q37854
ID Q37854 PRELIMINARY; PRT; 8 AA.
AC Q37854;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE COLIPHAGE GENE OF UNKNOWN FUNCTION (FRAGMENT).
OS Bacteriophage RL7.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;
OC Levivirus.
OX NCBI_TaxID=12026;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=73224987; PubMed=4352721;
RA Rensing U.F.E.;
```

RT "A sequence of seventy-three nucleotides from the coliphage R17  
 RT genome.";  
 RL Biochem. J. 131:593-604(1973).  
 DR EMBL; M24820; AAA72755.1; -.  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 969 MW; ECB45412C1E72726 CRC64;

Query Match 17.6%; Score 3; DB 9; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LLH 7

Db 2 LLH 4

RESULT 11

O95574 PRELIMINARY; PRT; 9 AA.

AC O95574;

DT 01-MAY-1999 (TReMBLrel. 10, Created)

DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)

DE SERUM ALBUMIN (FRAGMENT).

OS ALB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=FETAL LIVER;

RX MEDLINE=82081882; PubMed=6171778;

RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,

RA Najarian R.C., Seeburg P.H., Wion K.L.;

RT "The sequence of human serum albumin cDNA and its expression in E.

RT coli.";

RL Nucleic Acids Res. 9:6103-6114(1981).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=FETAL LIVER;

RX MEDLINE=82105994; PubMed=6275391;

RA Dugaiczky A., Law S.W., Dennison O.E.;

RT "Nucleotide sequence and the encoded amino acids of human serum

RT albumin mRNA";

RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=FETAL LIVER;

RX MEDLINE=85155492; PubMed=6085063;

RA Urano Y., Sakai M., Watanabe K., Tamaoki T.;

RT "Tandem arrangement of the albumin and alpha-fetoprotein genes in the

RT human genome.";

RL Gene 32:255-261(1984).

DR EMBL; L00133; AAC95394.1; -.

FT NON\_TER 1

SQ SEQUENCE 9 AA; 1186 MW; 92340B0400440681 CRC64;

Query Match 17.6%; Score 3; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 RER 17

Db 6 RER 8

RESULT 12

Q9UE26

ID Q9UE26 PRELIMINARY; PRT; 9 AA.

AC Q9UE26;

DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DE PEGI/MEST PROTEIN (FRAGMENT).  
 GN PEGI/MEST.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97336048; PubMed=9192843;

RA Riesewijk A.M., Hu L., Schulz U., Tariverdian G., Hoeglund P.,

RA Kere J., Ropers H.H., Kalscheuer V.M.;

RT "Monoallelic expression of human PEGI/MEST is paralleled by parent-

RT specific methylation in fetuses.";

RL Genomics 42:236-244(1997).

DR EMBL; Y10620; CAA71631.1; -.

FT NON\_TER 9

SQ SEQUENCE 9 AA; 1258 MW; C52D07340AB41412 CRC64;

Query Match 17.6%; Score 3; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LRR 15

Db 7 LRR 9

RESULT 13

O9GK05 PRELIMINARY; PRT; 9 AA.

AC O9GK05;

DT 01-MAR-2001 (TReMBLrel. 16, Created)

DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

DE ALPHA S2-CASEIN (FRAGMENT).

GN CSNLS2

OS Capra hircus (Goat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Capra.

OX NCBI\_TaxID=9925;

RN [1]

RP SEQUENCE FROM N.A.

RA Veltri C., Pilla F., Lagonigro R.;

RT "A new allele of goat alpha s2-casein.";

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ297312; CAC21706.1; -.

FT NON\_TER 1

FT NON\_TER 9

SQ SEQUENCE 9 AA; 1013 MW; 490FA5A1B1AEB723 CRC64;

Query Match 17.6%; Score 3; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 EKL 13

Db 1 EKL 3

RESULT 14

Q60192

ID Q60192 PRELIMINARY; PRT; 10 AA.

AC Q60192;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)

DE PHOSPHOFRUCTOKINASE (FRAGMENT).

GN PFK.  
OS Spiroplasma kunkelii.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Spiroplasmataceae; Spiroplasma.  
OX NCBI\_TaxID=47834;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=E275;  
RA Foissac X., Saillard C., Bove J.M.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U57659; AAB05469.1; -.  
FT NON\_TER 10  
SQ SEQUENCE 10 AA; 1103 MW; 58D39E2058705333 CRC64;

Query Match 17.6%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKI 3  
|||  
Db 3 KKI 5

RESULT 15  
Q60194  
ID Q60194 PRELIMINARY; PRT; 10 AA.  
AC Q60194;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
DE PHOSPHOFRUCTOKINASE (FRAGMENT).  
GN PFK.  
OS Spiroplasma phoeniceum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Spiroplasmataceae; Spiroplasma.  
OX NCBI\_TaxID=47835;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=P40;  
RA Foissac X., Saillard C., Bove J.M.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U57658; AAB05467.1; -.  
FT NON\_TER 10  
SQ SEQUENCE 10 AA; 1103 MW; 5C829E2058705333 CRC64;

Query Match 17.6%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKI 3  
|||  
Db 3 KKI 5

Search completed: September 15, 2001, 13:01:53  
Job time: 111 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:59:17 ; Search time 12.58 Seconds  
(without alignments)  
102.939 Million cell updates/sec

Title: us-09-389-000-2\_COPY\_134\_150

Perfect score: 17

Sequence: 1 KKISLLHSSKEKLRRR 17

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2758

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Listing first 45 summaries

Database : PIR.68.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4	23.5	6	2 B3932	Ig mu chain D regi
2	4	23.5	12	2 C49215	urease (EC 3.5.1.5
3	4	23.5	14	2 PC4382	dehydrin 4.5K poly
4	4	23.5	4	2 I40804	endoglucanase F-
5	3	17.6	5	2 I40698	biotin B - Citroba
6	3	17.6	6	2 A11490	pyruvate kinase (E
7	3	17.6	7	2 PS0254	18K protein 5507 -
8	3	17.6	8	2 PT0323	Ig heavy chain CRD
9	3	17.6	9	2 PT0231	Ig heavy chain CDR
10	3	17.6	10	2 S30348	clotting protein -
11	3	17.6	10	2 H61308	hemocyanin subunit
12	3	17.6	11	2 PT0081	protein QA300023 -
13	3	17.6	11	2 S19775	wound-induced prot
14	3	17.6	11	2 I54193	Rhesus blood group
15	3	17.6	11	2 PH0939	T-cell receptor be
16	3	17.6	12	2 S26558	T-cell receptor be
17	3	17.6	12	2 S01222	translation elonga
18	3	17.6	12	2 A58501	24K kidney and bla
19	3	17.6	12	2 D28551	hypothetical prote
20	3	17.6	12	2 T44420	hypothetical prote
21	3	17.6	12	2 A49637	MRC class II histo
22	3	17.6	12	2 A53524	ubiquinol-cytochr
23	3	17.6	13	2 S78519	ribosomal protein
24	3	17.6	13	2 JQ2309	hypothetical 1.6K
25	3	17.6	13	2 JQ2319	hypothetical 1.6K
26	3	17.6	13	2 B19434	probable sex-speci
27	3	17.6	13	2 S57571	T cell receptor al
28	3	17.6	13	2 PH0138	T-cell receptor be
29	3	17.6	13	2 PH1595	Ig H chain V-D-J r

30	3	17.6	13	4	I70076	glycophorin B/glyc
31	3	17.6	14	1	QMWAVV	mastoparan - yello
32	3	17.6	14	1	QMVHP2	mastoparan C - Eur
33	3	17.6	14	1	QMWAPP	polistes mastopara
34	3	17.6	14	2	B56819	PS I complex subun
35	3	17.6	14	2	C33098	223K exoantigen -
36	3	17.6	14	2	S45655	cathepsin L (EC 3.
37	3	17.6	14	2	PC1215	homeotic protein E
38	3	17.6	14	2	A61308	hemocyanin chain 2
39	3	17.6	14	2	B61308	hemocyanin chain 4
40	3	17.6	14	2	S41601	interferon alpha r
41	3	17.6	14	4	S00843	hypothetical prote
42	3	17.6	15	2	S61284	phosphoprotein, 80
43	3	17.6	15	2	A26228	spot 42 protein -
44	3	17.6	15	2	PA0053	protein QF200006 -
45	3	17.6	15	2	F44823	synaptosomal-assoc

ALIGNMENTS

RESULT 1

B3932

Ig mu chain D region (D23) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 16-Aug-1996

C:Accession: B3932

R:Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.

Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989

A:Title: Two murine natural polyreactive autoantibodies are encoded by nonmutated ger

A:Reference number: A3932; MUID:89282823

A:Accession: B3932

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-6 <BAC>

A:Cross-references: GB:M27107

C:Keywords: immunoglobulin

Query Match 23.5%; Score 4; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 EKLK 14

Db 1 EKLK 4

RESULT 2

C49215

urease (EC 3.5.1.5) large subunit UreB - Helicobacter felis (fragment)

C:Species: Helicobacter felis

C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 31-Oct-1997

C:Accession: C49215

R:Turbett, G.R.; Hoj, P.B.; Horne, R.; Mee, B.J.

Infect. Immun. 60, 5259-5266, 1992

A:Title: Purification and characterization of the urease enzymes of Helicobacter spec

A:Reference number: A49215; MUID:93084378

A:Accession: C49215

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-12 <TUR>

A:Experimental source: ATCC 49179

A:Note: sequence extracted from NCBI backbone (NCBIP:119484)

C:Superfamily: urease 62K chain; urease 62K chain homology

C:Keywords: hydrolase

Query Match 23.5%; Score 4; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKIS 4

Db 2 KKIS 5  
||||

RESULT 3  
PC4382  
dehydrin 4.5K polypeptide - Soybean (fragment)  
N:Alternate names: acid soluble 26K protein  
C:Species: Glycine max cv. Nattoyoryu (soybean)  
C:Date: 06-Nov-1997 #sequence\_revision 06-Nov-1997 #text\_change 07-May-1999  
C:Accession: PC4382  
R:Momma, M.; Haraguchi, K.; Saito, M.; Chikuni, K.; Harada, K.  
Biosci. Biotechnol. Biochem. 61, 1286-1291, 1997  
A:Title: Purification and characterization of the acid soluble 26-kDa polypeptide from soybean  
A:Reference number: PC4380; MUID:97446521  
A:Accession: PC4382  
A:Molecule type: protein  
A:Residues: 1-14 <NOM>  
A:Experimental source: seed  
C:Comment: This protein is rich in hydrophilic amino acid, and highly heat stable.

Query Match 23.5%; Score 4; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KEKL 13  
||||  
Db 5 KEKL 8

RESULT 4  
I40804  
endoglucanase F - Clostridium thermocellum (fragment)  
C:Species: Clostridium thermocellum  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 16-Aug-1996  
C:Accession: I40804  
R:Mishra, S.; Beguin, P.; Aubert, J.  
J. Bacteriol. 173, 80-85, 1991  
A:Title: Transcription of clostridium thermocellum endoglucanase genes celf and celd.  
A:Reference number: I40804; MUID:91100322  
A:Accession: I40804  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-4 <RES>  
A:Cross-references: GB:M64363; NID:gl44771  
C:Genetics:  
A:Gene: celf  
A:Start codon: TTG

Query Match 17.6%; Score 3; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKI 3  
|||  
Db 2 KKI 4

RESULT 5  
I40698  
biotin B - Citrobacter freundii (fragment)  
C:Species: Citrobacter freundii  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 12-Aug-1996  
C:Accession: I40698  
R:Shuan, D.; Campbell, A.  
Gene 67, 203-211, 1988  
A:Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citrobacter freundii  
A:Reference number: I40697; MUID:85006280  
A:Accession: I40698  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA

A:Residues: 1-5 <RES>  
A:Cross-references: GB:M21922; NID:gl144434

Query Match 17.6%; Score 3; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 HSS 9  
|||  
Db 3 HSS 5

RESULT 6  
A11490  
pyruvate kinase (EC 2.7.1.40) - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 03-Mar-1995  
C:Accession: A11490  
R:Hjeltnquist, G.; Andersson, J.; Edlund, B.; Engstrom, L.  
Biochem. Biophys. Res. Commun. 61, 559-563, 1974  
A:Title: Amino acid sequence of a (32-P)phosphopeptide from pig liver pyruvate kinase  
A:Reference number: A11490; MUID:75127438  
A:Accession: A11490  
A:Molecule type: protein  
A:Residues: 1-6 <HJE>  
A:Experimental source: liver  
C:Keywords: glycolysis; phosphotransferase

Query Match 17.6%; Score 3; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRR 15  
|||  
Db 1 LRR 3

RESULT 7  
PS0254  
18K protein 5507 - rice (strain Nihonbare) (fragment)  
C:Species: Oryza sativa (rice)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 23-Mar-1995  
C:Accession: PS0254  
R:Tsugita, A.  
submitted to JIPID, April 1993  
A:Reference number: PS0206  
A:Accession: PS0254  
A:Molecule type: protein  
A:Residues: 1-7 <FSU>  
A:Experimental source: leaf, chloroplast, strain Nihonbare  
A:Note: molecular weight 18K, pI 4.4

Query Match 17.6%; Score 3; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KEK 12  
|||  
Db 5 KEK 7

RESULT 8  
PT0323  
Ig heavy chain CRD3 region (clone J2-106B) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0323  
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an

A:Reference number: PT0222; MUID:91108337  
A:Accession: PT0323  
A:Molecule type: DNA  
A:Residues: 1-8 <YAM>  
A:Experimental source: B lymphocyte  
C:Keywords: heterotetramer; immunoglobulin

Query Match 17.6%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 RRE 16  
|||  
Db 6 RRE 8

## RESULT 9

PT0231  
Ig heavy chain CDR3 region (clone 1-118B) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
C:Accession: PT0231  
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991  
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and  
A:Reference number: PT0222; MUID:91108337  
A:Accession: PT0231  
A:Molecule type: DNA  
A:Residues: 1-9 <YAM>  
A:Experimental source: B lymphocyte  
C:Keywords: heterotetramer; immunoglobulin

Query Match 17.6%; Score 3; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 HSS 9  
|||  
Db 3 HSS 5

## RESULT 10

S30348  
clotting protein - signal crayfish  
C:Species: Pacifastacus leniusculus (signal crayfish)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 17-Mar-1999  
C:Accession: S30348  
R:Kopacek, P.; Hall, M.; Soederhaell, K.  
Eur. J. Biochem. 213, 591-597, 1993  
A:Title: Characterization of a clotting protein, isolated from plasma of the freshwater  
A:Reference number: S30348; MUID:93238739  
A:Accession: S30348  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <KOP>

Query Match 17.6%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.9e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LHS 8  
|||  
Db 1 LHS 3

## RESULT 11

H61308  
hemocyanin subunit IV - Atlantic horseshoe crab (fragment)  
C:Species: Limulus polyphemus (Atlantic horseshoe crab)  
C:Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 07-May-1999

C:Accession: H61308  
R:Jolles, J.; Jolles, P.; Lamy, J.; Lamy, J.  
FEBS Lett. 106, 289-291, 1979  
A:Title: Structural characterization of seven different subunits in Androctonus austr  
A:Reference number: A61308; MUID:80047238  
A:Accession: H61308  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <JOL>

Query Match 17.6%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.9e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KEK 12  
|||  
Db 3 KEK 5

## RESULT 12

PT0081  
protein QA300023 - Arabidopsis thaliana (fragment)  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 07-Feb-1996 #sequence\_revision 19-Apr-1996 #text\_change 24-Nov-1999  
C:Accession: PT0081  
R:Tsugita, A.; Kamo, M.; Kawakami, M.; Ohki, Y.  
submitted to JIPID, December 1995  
A:Description: Two dimensional electrophoresis of plant proteins and standardization  
A:Reference number: PN0173  
A:Accession: PT0081  
A:Molecule type: protein  
A:Residues: 1-11 <TSU>  
A:Experimental source: Leaf  
C:Keywords: acetylated amino end  
P:1/Modified site: acetylated amino end (Ser) #status experimental

Query Match 17.6%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 4.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 SKE 11  
|||  
Db 1 SKE 3

## RESULT 13

S19775  
wound-induced protein - tomato (fragment)  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Sep-1997  
C:Accession: S19775  
R:Parsons, B.L.  
submitted to the EMBL Data Library, May 1991  
A:Reference number: S19773  
A:Accession: S19775  
A:Molecule type: mRNA  
A:Residues: 1-11 <PAR>  
A:Cross-references: EMBL:X59884; NID:g19323; PID:g19324

Query Match 17.6%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 4.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSK 10  
|||  
Db 3 SSK 5

## RESULT 14

154193

Rhesus blood group CcEe protein - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: I54193  
R:Cherif-Zahar, B.; Le Van Kim, C.; Rouillac, C.; Raynal, V.; Cartron, J.P.; Collin, Y.  
Genomics 19, 68-74, 1994  
A:Title: Organization of the gene (RHCE) encoding the human blood group RhCcEe antigens  
A:Reference number: I54193; MUID:94245182  
A:Accession: I54193  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-11 <RES>  
A:Cross-references: GB:S70456; NID:g546795; PIDN:AAD14061.1; PID:g4261761  
C:Genetics:  
A:Gene: GDB:RHCE  
A:Cross-references: GDB:229957; OMIM:111700  
A:Map position: lp36.2-lp34

Query Match 17.6%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 4.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSK 10  
|||  
Db 2 SSK 4

RESULT 15  
PH0939  
T-cell receptor beta chain V-D-J region (clone 10) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
C:Accession: PH0939  
R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
J. Exp. Med. 174, 1467-1476, 1991  
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic  
A:Reference number: PH0891; MUID:92078857  
A:Accession: PH0939  
A:Molecule type: mRNA  
A:Residues: 1-11 <GOL>  
A:Experimental source: complete Freund's adjuvant-immunized lymph node  
C:Keywords: T-cell receptor

Query Match 17.6%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 4.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSK 10  
|||  
Db 3 SSK 5

Search completed: September 15, 2001, 13:01:06  
Job time: 109 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:58:57 ; Search time 11.44 Seconds  
(without alignments)  
30.598 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_150

Perfect score: 17

Sequence: 1 KKISLLHSSKEKLRRR 17

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 91644

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgnl\_7/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgnl\_7/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgnl\_7/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgnl\_7/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgnl\_7/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgnl\_7/ptodata/1/iaa/backfiles1.pep.\*

Pred. No: is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5	29.4	6	5	PCT-US93-12679-5
2	5	29.4	16	1	US-08-438-123-1
3	4	23.5	4	2	US-08-441-871-52
4	4	23.5	5	1	US-08-704-170-14
5	4	23.5	5	2	US-08-350-260A-415
6	4	23.5	5	5	PCT-US94-02631-14
7	4	23.5	6	1	US-08-704-170-4
8	4	23.5	6	2	US-08-482-228-160
9	4	23.5	6	3	US-08-482-228-160
10	4	23.5	6	5	PCT-US94-02631-4
11	4	23.5	7	3	US-08-834-314-13
12	4	23.5	7	3	US-09-079-447-2
13	4	23.5	7	4	US-09-025-819-5
14	4	23.5	9	3	US-08-159-339A-135
15	4	23.5	9	3	US-08-159-339A-252
16	4	23.5	10	1	US-08-378-761A-46
17	4	23.5	10	1	US-08-485-286-46
18	4	23.5	10	1	US-08-386-579-9
19	4	23.5	10	1	US-08-386-579-11
20	4	23.5	10	1	US-08-139-508-52
21	4	23.5	10	1	US-08-199-508-53
22	4	23.5	10	2	US-08-726-306A-1
23	4	23.5	10	2	US-08-665-055-3
24	4	23.5	10	2	US-08-665-055-5
25	4	23.5	10	3	US-08-342-924-8
26	4	23.5	10	3	US-08-342-924-10
27	4	23.5	10	4	US-09-326-840-3

28	4	23.5	10	4	US-09-326-840-5	Sequence 5, Appli
29	4	23.5	10	4	US-09-177-249-235	Sequence 235, App
30	4	23.5	10	4	US-08-665-057-3	Sequence 3, Appli
31	4	23.5	10	4	US-08-665-057-5	Sequence 5, Appli
32	4	23.5	10	5	PCT-US91-08328-24	Sequence 24, Appli
33	4	23.5	10	5	PCT-US96-01807-9	Sequence 9, Appli
34	4	23.5	10	5	PCT-US96-01807-11	Sequence 11, Appli
35	4	23.5	10	6	5210075-7	Patent No. 5210075
36	4	23.5	10	6	5248606-35	Patent No. 5248606
37	4	23.5	11	2	US-08-342-930-5	Sequence 5, Appli
38	4	23.5	11	4	US-08-392-542-13	Sequence 13, Appli
39	4	23.5	11	4	US-08-894-327-13	Sequence 13, Appli
40	4	23.5	11	5	PCT-US91-08328-14	Sequence 14, Appli
41	4	23.5	11	5	PCT-US91-08328-16	Sequence 16, Appli
42	4	23.5	11	5	PCT-US91-08328-18	Sequence 18, Appli
43	4	23.5	12	1	US-07-778-233B-13	Sequence 13, Appli
44	4	23.5	12	1	US-07-963-321-13	Sequence 13, Appli
45	4	23.5	12	1	US-08-290-641-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1  
PCT-US93-12679-5  
; Sequence 5, Application PC/TUS9312679  
; GENERAL INFORMATION:  
; APPLICANT: Rath, Matthias  
; TITLE OF INVENTION: SYNTHETIC OLIGOPEPTIDES ANALOGOUS TO  
; TITLE OF INVENTION: PROTEIN SIGNAL SEQUENCES -METHODS OF IDENTIFICATION AND  
; TITLE OF INVENTION: METHODS OF USE  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHELDON & MAK  
; STREET: 401 Florence Street  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/12679  
; FILING DATE: 30-DEC-1993  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cranfill, Raymond B  
; REGISTRATION NUMBER: 32,845  
; REFERENCE/DOCKET NUMBER: RATH-10016PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-322-5333  
; TELEFAX: 415-322-5499  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
PCT-US93-12679-5

Query Match 29.4%; Score 5; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSKEK 12

Db 2 SSKEK 6

RESULT 2  
US-08-438-123-1  
; Sequence 1, Application US/08438123  
; Patent No. 5552293  
; GENERAL INFORMATION:  
; APPLICANT: Lindholm et al  
; TITLE OF INVENTION: TUMOR ANTIGEN SPECIFIC ANTIBODY  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lowe, Price, Leblanc & Becker  
; STREET: Suite 300, 99 Canal Center Plaza  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22314  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: DOS Text File  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/438,123  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/906,350  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: J.G. Mullins  
; REGISTRATION NUMBER: 33073  
; REFERENCE/DOCKET NUMBER: 149-011  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703 684 1111  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Polypeptide  
US-08-438-123-1

Query Match 29.4%; Score 5; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 4 SLLHS 8  
Db 5 SLLHS 9

RESULT 3  
US-08-441-871-52  
; Sequence 52, Application US/08441871  
; Patent No. 5846765  
; GENERAL INFORMATION:  
; APPLICANT: Matthews, David J.  
; APPLICANT: Wells, James A.  
; APPLICANT: Zoller, Mark J.  
; TITLE OF INVENTION: Identification of No. 5846765el Substrates  
; NUMBER OF SEQUENCES: 152  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/441,871  
; FILING DATE: 16-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/418928  
; FILING DATE: 05-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/161692  
; FILING DATE: 03-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/864452  
; FILING DATE: 06-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/09133  
; FILING DATE: 03-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/743614  
; FILING DATE: 09-AUG-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715300  
; FILING DATE: 14-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/683400  
; FILING DATE: 10-APR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/621667  
; FILING DATE: 03-DEC-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winter, Daryl B.  
; REGISTRATION NUMBER: 32,637  
; REFERENCE/DOCKET NUMBER: 645PSC2D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1249  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 52:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-441-871-52  
Query Match 23.5%; Score 4; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 8 SSKE 11  
Db 1 SSKE 4  
RESULT 4  
US-08-704-170-14  
; Sequence 14, Application US/08704170  
; Patent No. 5707626  
; GENERAL INFORMATION:  
; APPLICANT: Douvas, Angeline  
; APPLICANT: Takehana, Yoshi  
; APPLICANT: Ehresmann, Glenn  
; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
; IMMUNOINFECTION: IMMUNOINFECTION CLUSTER VIRUS INFECTIONS  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Robbins, Berliner & Carson  
; STREET: 201 No. 5707626th Figueroa Street, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90012  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/704.170  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-704-170-14

Query Match 23.5%; Score 4; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 RRER 17  
||||  
Db 1 RRER 4

RESULT 5  
US-08-350-260A-415  
Sequence 415, Application US/08350260A  
Patent No. 5962255  
GENERAL INFORMATION:  
APPLICANT: Winter, Gregory Paul  
APPLICANT: Griffiths, Andrew David  
APPLICANT: Williams, Samuel Cameron  
APPLICANT: Waterhouse, Peter  
APPLICANT: Nissim, Ahuva  
APPLICANT: Johnson, Kevin Stuart  
APPLICANT: Smith, Andrew John Hammond  
TITLE OF INVENTION: Methods for producing members of specific  
TITLE OF INVENTION: binding pairs  
NUMBER OF SEQUENCES: 602  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David W. Clough  
STREET: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/350,260A  
FILING DATE: 05-DEC-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9110549.4  
FILING DATE: 15-MAY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9206318.9

FILING DATE: 24-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB91/01134  
FILING DATE: 10-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB92/00883  
FILING DATE: 15-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/00605  
FILING DATE: 24-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/150,002  
FILING DATE: 31-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/307,619  
FILING DATE: 16-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 28111/32372  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
INFORMATION FOR SEQ ID NO: 415:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-350-260A-415

Query Match 23.5%; Score 4; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRRE 16  
||||  
Db 1 LRRE 4

RESULT 6  
PCT-US94-02631-14  
Sequence 14, Application PC/TUS9402631  
GENERAL INFORMATION:  
APPLICANT: Douvas, Angelina  
APPLICANT: Takehana, Yoshi  
APPLICANT: Ehresmann, Glenn  
TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
TITLE OF INVENTION: IMMUNOINFECTIVE CLUSTER VIRUS INFECTIONS  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Robbins, Berliner & Carson  
STREET: 201 North Figueroa Street, Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/02631  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US94-02631-14

Query Match 23.5%; Score 4; DB 5; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 RRER 17  
||||

Db 1 RRER 4

RESULT 7  
US-08-704-170-4  
; Sequence 4, Application US/08704170  
; Patent No. 5707626  
; GENERAL INFORMATION:  
; APPLICANT: Douvas, Angeline  
; APPLICANT: Takehana, Yoshi  
; APPLICANT: Ehresmann, Glenn  
; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Robbins, Berliner & Carson  
; STREET: 201 No. 5707626th Figueroa Street, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90012

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/704,170  
FILING DATE:

CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-704-170-4

Query Match 23.5%; Score 4; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 RRER 17  
||||

Db 1 RRER 4

RESULT 8  
US-08-482-228-160  
; Sequence 160, Application US/08482228  
; Patent No. 5968753  
; GENERAL INFORMATION:  
; APPLICANT: Tseng-Law, Janet  
; APPLICANT: Kobori, Joan A.  
; APPLICANT: Al-Abdaly, Fahad A.  
; APPLICANT: Guillermo, Roy  
; APPLICANT: Helgerson, Sam L.  
; APPLICANT: Deans, Robert J.  
; TITLE OF INVENTION: POSITIVE AND POSITIVE/NEGATIVE CELL  
; TITLE OF INVENTION: SELECTION MEDIATED BY PEPTIDE RELEASE  
; NUMBER OF SEQUENCES: 215  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Janice Guthrie, Ph.D.  
; STREET: P.O. Box 15210  
; CITY: Irvine  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92713-5210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,228  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Guthrie, Janice  
REGISTRATION NUMBER: 35,170  
REFERENCE/DOCKET NUMBER: IT-4630CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (714) 440-5353  
TELEFAX: (714) 553-1952  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-482-228-160

Query Match 23.5%; Score 4; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLH 7  
||||

Db 3 SLLH 6

RESULT 9  
US-08-482-528-160  
; Sequence 160, Application US/08482528  
; Patent No. 6017719  
; GENERAL INFORMATION:  
; APPLICANT: Tseng-Law, Janet  
; APPLICANT: Kobori, Joan A.  
; APPLICANT: Al-Abdaly, Fahad A.  
; APPLICANT: Guillermo, Roy  
; APPLICANT: Helgerson, Sam L.  
; APPLICANT: Deans, Robert J.  
; TITLE OF INVENTION: POSITIVE AND POSITIVE/NEGATIVE CELL  
; TITLE OF INVENTION: SELECTION MEDIATED BY PEPTIDE RELEASE  
; NUMBER OF SEQUENCES: 215  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Janice Guthrie, Ph.D.  
STREET: P.O. Box 15210  
CITY: Irvine  
STATE: California  
COUNTRY: USA  
ZIP: 92713-5210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,528  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Guthrie, Janice  
REGISTRATION NUMBER: 35,170  
REFERENCE/DOCKET NUMBER: IT-4630CIP4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (714) 440-5353  
TELEFAX: (714) 553-1952  
INFORMATION FOR SEQ ID NO: 160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-482-528-160

Query Match 23.5%; Score 4; DB 3; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLH 7  
DB 3 SLH 6

RESULT 10  
PCT-US94-02631-4  
; Sequence 4, Application PC/TUS9402631  
; GENERAL INFORMATION:  
; APPLICANT: Douvas, Angelina  
; APPLICANT: Takehana, Yoshi  
; APPLICANT: Ehresmann, Glenn  
; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
; IMMUNOINFECTION: IMMUNOINFECTION CLUSTER VIRUS INFECTIONS  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Robbins, Berliner & Carson  
; STREET: 201 North Figueroa Street, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90012  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/02631  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/029,850  
; FILING DATE: 11-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spitals, John P.  
; REGISTRATION NUMBER: 29,215  
; REFERENCE/DOCKET NUMBER: 1920-331

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US94-02631-4

Query Match 23.5%; Score 4; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 RER 17  
DB 1 RER 4

RESULT 11  
US-08-834-314-13  
; Sequence 13, Application US/08834314  
; Patent No. 6080911  
; GENERAL INFORMATION:  
; APPLICANT: Kopchick, John J.  
; APPLICANT: Zhon, Yihua  
; TITLE OF INVENTION: Animal Models of Growth Hormone  
; TITLE OF INVENTION: Insensitivity  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/834,314  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ingolia, Diane E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: OHU-02577  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: protein  
US-08-834-314-13

Query Match 23.5%; Score 4; DB 3; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKS 4  
DB 4 KKS 7

## RESULT 12

US-09-079-447-2  
; Sequence 2, Application US/09079447  
; Patent No. 6083486  
; GENERAL INFORMATION:  
; APPLICANT: Weissleder, Ralph  
; APPLICANT: Tung, Ching  
; APPLICANT: Mahmood, Umar  
; APPLICANT: Josephson, Lee  
; APPLICANT: Bogdanov, Alexei  
; TITLE OF INVENTION: INTRAMOLECULARLY-QUENCHED  
; TITLE OF INVENTION: NEAR INFRARED FLUORESCENT PROBES  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA

; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FASTSEQ for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/079,447  
; FILING DATE: 14-MAY-1998  
; PRIOR APPLICATION NUMBER:  
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:  
; NAME: Creason, Gary L.  
; REGISTRATION NUMBER: 34,310  
; REFERENCE/DOCKET NUMBER: 08472/758001  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-079-447-2

Query Match 23.5% Score 4; DB 3; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 HSSK 10  
| | | |  
DB 1 HSSK 4

## RESULT 13

US-09-025-819-5  
; Sequence 5, Application US/09025819  
; Patent No. 6225097  
; GENERAL INFORMATION:  
; APPLICANT: Obata, Shusei  
; APPLICANT: Nishino, Tokuzo  
; APPLICANT: Koyama, Tanetoshi  
; APPLICANT: Sato, Yoshihiro  
; TITLE OF INVENTION: DECAPRENYL DIPHOSPHATE SYNTHETASE GENE  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: KENYON & KENYON  
; STREET: 1500 K Street, N.W.  
; CITY: Washington  
; STATE: DC

; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/025,819  
; FILING DATE: 19-FEB-1998  
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 251675  
; FILING DATE: 17-SEP-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Khalilian, Hourii  
; REGISTRATION NUMBER: 39,546  
; REFERENCE/DOCKET NUMBER: 10235/2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-220-4200  
; TELEFAX: 202-220-4201  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-025-819-5

Query Match 23.5% Score 4; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SLLH 7  
| | | |  
DB 2 SLLH 5

## RESULT 14

US-08-159-339A-135  
; Sequence 135, Application US/08159339A  
; Patent No. 6037135  
; GENERAL INFORMATION:  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Celis, Esteban  
; TITLE OF INVENTION: HLA Binding peptides and Their  
; TITLE OF INVENTION: Uses  
; NUMBER OF SEQUENCES: 1254  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA

; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/159,339A  
; FILING DATE: 29-NOV-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/926,666  
; FILING DATE: 07-AUG-1992  
; APPLICATION NUMBER: US 08/027,746  
; FILING DATE: 05-MAR-1993

; APPLICATION NUMBER: US 08/103,396  
; FILING DATE: 06-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauver  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 018623-005030US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 135:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-159-339A-135

Query Match 23.5%; Score 4; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LRRE 16  
||||  
Db 4 LRRE 7

RESULT 15  
US-08-159-339A-252  
; Sequence 252, Application US/08159339A  
; Patent No. 6037135  
; GENERAL INFORMATION:  
; APPLICANT: Kubo, Ralph T.  
; APPLICANT: Grey, Howard M.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Cellis, Esteban  
; TITLE OF INVENTION: HLA Binding peptides and Their  
; TITLE OF INVENTION: Uses  
; NUMBER OF SEQUENCES: 1254  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/159,339A  
; FILING DATE: 29-NOV-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/926,666  
; FILING DATE: 07-AUG-1992  
; APPLICATION NUMBER: US 08/027,746  
; FILING DATE: 05-MAR-1993  
; APPLICATION NUMBER: US 08/103,396  
; FILING DATE: 06-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen Lauver  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 018623-005030US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 252:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-159-339A-252

Query Match 23.5%; Score 4; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LRRE 16  
||||  
Db 2 LRRE 5

Search completed: September 15, 2001, 13:00:45  
Job time: 108 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:58:37 ; Search time 17.59 Seconds  
(without alignments)  
58.591 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_150  
Perfect score: 17  
Sequence: 1 KKTSLHSSKEKLRRR 17

Scoring table:  
OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 138608

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Listing first 45 summaries

- Database : A\_Geneseq\_0601.\*
- 1: /cgnl\_9/gcgdata/geneseq/geneseq/AA1980.DAT.\*
  - 2: /cgnl\_9/gcgdata/geneseq/geneseq/AA1981.DAT.\*
  - 3: /cgnl\_9/gcgdata/geneseq/geneseq/AA1982.DAT.\*
  - 4: /cgnl\_9/gcgdata/geneseq/geneseq/AA1983.DAT.\*
  - 5: /cgnl\_9/gcgdata/geneseq/geneseq/AA1984.DAT.\*
  - 6: /cgnl\_9/gcgdata/geneseq/geneseq/AA1985.DAT.\*
  - 7: /cgnl\_9/gcgdata/geneseq/geneseq/AA1986.DAT.\*
  - 8: /cgnl\_9/gcgdata/geneseq/geneseq/AA1987.DAT.\*
  - 9: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT.\*
  - 10: /cgnl\_9/gcgdata/geneseq/geneseq/AA1989.DAT.\*
  - 11: /cgnl\_9/gcgdata/geneseq/geneseq/AA1990.DAT.\*
  - 12: /cgnl\_9/gcgdata/geneseq/geneseq/AA1991.DAT.\*
  - 13: /cgnl\_9/gcgdata/geneseq/geneseq/AA1992.DAT.\*
  - 14: /cgnl\_9/gcgdata/geneseq/geneseq/AA1993.DAT.\*
  - 15: /cgnl\_9/gcgdata/geneseq/geneseq/AA1994.DAT.\*
  - 16: /cgnl\_9/gcgdata/geneseq/geneseq/AA1995.DAT.\*
  - 17: /cgnl\_9/gcgdata/geneseq/geneseq/AA1996.DAT.\*
  - 18: /cgnl\_9/gcgdata/geneseq/geneseq/AA1997.DAT.\*
  - 19: /cgnl\_9/gcgdata/geneseq/geneseq/AA1998.DAT.\*
  - 20: /cgnl\_9/gcgdata/geneseq/geneseq/AA1999.DAT.\*
  - 21: /cgnl\_9/gcgdata/geneseq/geneseq/AA2000.DAT.\*
  - 22: /cgnl\_9/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	11	64.7	15	21	AA199270
2	5	29.4	6	4	AA199270
3	5	29.4	6	15	AA199270
4	5	29.4	9	15	AA199270
5	5	29.4	14	20	AA199270
6	5	29.4	16	14	AA199270
7	5	29.4	16	20	AA199270
8	5	29.4	16	21	AA199270
9	4	23.5	4	20	AA199270
10	4	23.5	5	15	AA199270
11	4	23.5	5	17	AA199270

12	4	23.5	5	19	AA199270
13	4	23.5	6	15	AA199270
14	4	23.5	6	16	AA199270
15	4	23.5	6	20	AA199270
16	4	23.5	6	20	AA199270
17	4	23.5	6	21	AA199270
18	4	23.5	6	21	AA199270
19	4	23.5	6	22	AA199270
20	4	23.5	7	12	AA199270
21	4	23.5	7	20	AA199270
22	4	23.5	7	21	AA199270
23	4	23.5	7	21	AA199270
24	4	23.5	7	21	AA199270
25	4	23.5	7	21	AA199270
26	4	23.5	7	21	AA199270
27	4	23.5	7	21	AA199270
28	4	23.5	7	22	AA199270
29	4	23.5	8	8	AA199270
30	4	23.5	8	15	AA199270
31	4	23.5	8	19	AA199270
32	4	23.5	8	19	AA199270
33	4	23.5	8	21	AA199270
34	4	23.5	8	21	AA199270
35	4	23.5	8	21	AA199270
36	4	23.5	8	21	AA199270
37	4	23.5	8	21	AA199270
38	4	23.5	9	14	AA199270
39	4	23.5	9	14	AA199270
40	4	23.5	9	14	AA199270
41	4	23.5	9	18	AA199270
42	4	23.5	9	18	AA199270
43	4	23.5	9	19	AA199270
44	4	23.5	9	19	AA199270
45	4	23.5	9	20	AA199270

ALIGNMENTS

RESULT 1  
AA199270  
ID AA199270 standard; Peptide; 15 AA.  
XX  
XX AA199270:  
AC  
XX  
DT 03-JUL-2000 (first entry)  
XX  
DE PHE19X peptide used to raise antibody.  
XX  
KW PHE19X; human; testis-specific; transcription factor;  
KW prostate cancer; bladder cancer; Ovary cancer; testicular cancer;  
KW therapy; diagnosis; vaccine; antibody.  
XX  
OS Homo sapiens.  
XX  
PN WO200012709-A2.  
XX  
PD 09-MAR-2000.  
XX  
PF 31-AUG-1999; 99WO-US20137.  
XX  
PR 31-AUG-1998; 98US-0098610.  
PR 31-OCT-1998; 98US-0106524.  
XX  
XX (UROC-) UROGENESYS INC.  
XX (AFAR/) AFAR D E.  
XX (HUBE/) HUBERT R S.  
XX (RAIT/) RAITANO A B.  
PI Afar DE, Hubert RS, Raitano AB;  
XX WPI; 2000-237872/20.  
DR  
XX

Human reflex tears  
Hydrophilic motif  
Farnesyl synthetase  
ATCC HB 11885 mono  
Peptide #1 cleavag  
PSA substrate pept  
Human haematopoiet  
Peptide used in ta  
ID1 plasmidium sur  
Peptide #2 cleavag  
Peptide useful for  
Rat neuronal T-typ  
Rat nervous system  
PSA substrate pept  
Cathepsin D-cleava  
Human brain T calc  
Sequence of new an  
Dynorphin-like pol  
Mouse growth hormo  
Helicobacter pylor  
Peptide associated  
Peptide associated  
C-terminal fragmen  
Glutamic acid deca  
Human hyaluronic a  
MHC Class I allele  
MHC Class I allele  
MHC Class I allele  
HPV16 E6 peptide (  
HPV16 E6-encoded C  
Peptide from HPV 1  
Peptide from HPV 1  
Amino acid sequenc

PT Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -

XX Example 5; Page 36; 62pp; English.

XX The present sequence is that of a peptide derived from human  
 CC PHELIX (see AAY79269), a novel transcription factor that is  
 CC normally expressed only in testis tissue, but which is up-regulated  
 CC in prostate and some other cancers. The peptide was conjugated to  
 CC keyhole limpet haemocyanin and used to raise polyclonal antiserum  
 CC in rabbit. The antiserum demonstrated specificity for PHELIX and  
 CC may therefore be useful for assessing the expression of PHELIX in  
 CC patient samples.

XX Sequence 15 AA;

Query Match 64.7%; Score 11; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.00026;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 HSSKEKLRRER 17  
 Db 1 hsskeklrrer 11  
 |||||

RESULT 2

AAP30295  
 ID AAP30295 standard; Protein; 6 AA.

XX AC AAP30295;

XX DT 20-APR-1992 (first entry)

XX DE Sequence of AAs 26-31 of the E. coli heat stable toxin  
 DE which determine an H-epitope.

XX KW Synthetic vaccine; antigen; allergen; immunological response;  
 KW antibody.

XX OS Escherichia coli.

XX PN EP3851-A.

XX PD 16-NOV-1983.

XX PF 11-MAR-1983; 83EP-0102392.

XX PR 15-MAR-1982; 82US-0358150.

XX PR 28-JAN-1983; 83US-0461802.

XX PR 12-JUN-1981; 81US-0272855.

XX PR 09-JAN-1981; 81US-0223558.

XX PR 16-DEC-1986; 86US-0942562.

XX PA (NYBL-) NEW YORK BLOOD CENT.

XX PI Hopp TP;

XX WPI; 1983-822049/47.

XX Synthetic vaccine - contains peptide residue coupled to higher  
 alkyl or alkenyl Gps. and with 6 amino acids in residue

XX Claim 50; Page 46; 54pp; English.

XX The inventors claim a synthetic vaccine which comprises a peptide  
 CC residue coupled to an alkyl or alkenyl gp. having at least 12C, or  
 CC other lipophilic substance. The residue contains a sequence of 6 AAs  
 CC corresp. to the SQ of such AAs in a protein antigen or allergen where  
 CC the greatest local average hydrophilicity is found. Pref. the AAs in  
 CC the peptide do not exceed 50 residues, and they esp. contain 12-18  
 CC residues. The alkyl or alkenyl gp. pref. contains 12-24C and it is

CC pref. coupled to the terminal amino gp. of the residue opt. via a CO  
 CC gp. If a lipophilic substance is used, it is pref. palmitic, stearic,  
 CC behenic, oleic or mycolic acid.

XX SQ Sequence 6 AA;

Query Match 29.4%; Score 5; DB 4; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSKEK 12  
 Db 2 sskek 6  
 |||||

RESULT 3

AAR59948  
 ID AAR59948 standard; peptide; 6 AA.

XX AC AAR59948;

XX DT 14-FEB-1995 (first entry)

XX DE Peptide signal sequence for treating E. coli infections.

XX KW Therapeutic; metabolic interactions; PSS; analogues.

XX OS Synthetic.

XX PN WO9416328-A.

XX PD 21-JUL-1994.

XX PF 30-DEC-1993; 93WO-US12679.

XX PR 30-DEC-1992; 92US-0997727.

XX PA (RATH/) RATH M.

XX PI Rath M;

XX WPI; 1994-249399/30.

XX Identifying peptide signal sequences in a protein - and use of  
 PT their synthetic analogues for treating or preventing, e.g. of  
 PT cardiovascular and auto-immune disease, infections and cancer.

XX Claim 18; Page 10; 28pp; English.

XX The sequence is that of a peptide signal sequence which can be used  
 CC to treat E. coli infections.  
 CC See also AAR59944-83.

XX SQ Sequence 6 AA;

Query Match 29.4%; Score 5; DB 15; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSKEK 12  
 Db 2 sskek 6  
 |||||

RESULT 4

AAR61019  
 ID AAR61019 standard; peptide; 9 AA.

XX AC AAR61019;

XX DT 23-APR-1995 (first entry)

XX Dynorphin-like polypeptide.  
 XX Dynorphin; opioid analgesic; stable; stability; intravenous.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FT Modified-site 1 /note= "MeTyr"  
 FT Misc-difference 8  
 FT /note= "D-Glu"  
 FT Modified-site 9  
 FT /note= "Arg-NH2"  
 XX EP614913-A.  
 XX 14-SEP-1994.  
 XX 08-NOV-1985; 94EP-0107769.  
 XX 09-NOV-1984; 84JP-0236076.  
 XX (EISA ) EISAI CO LTD.  
 XX Arakawa Y, Araki S, Ikeda M, Kaneko T, Nakazawa T;  
 PI Tachibana S, Tsuchiya Y, Yamatsu K, Yoshino H;  
 XX WPI; 1994-281157/35.  
 XX New dynorphin polypeptide derivs. for use as analgesics - contg.  
 PT a D-amino acid or an N-alkyl deriv. of a D- or L-amino acid for  
 PT in vivo stability.  
 XX Example 26; Page 27; 47pp; English.  
 XX The patent discloses dynorphin-like polypeptides which are shorter  
 CC than dynorphin (7 to 9 amino acids) and which contain either a D-  
 CC amino acid and/or an N-alkyl deriv. of an L- or D-amino acid to  
 CC provide in-vivo stability upon intravenous administration. The  
 CC peptides have a much greater analgesic effect than dynorphin  
 CC due to their greater stability in the blood (e.g. IC50 values can be  
 CC as low as 0.04 nM when tested by the rabbit vas deferens method; cf.  
 CC 17.4 nM for dynorphin).  
 CC The present peptide is a specific example of the peptides.  
 XX Sequence 9 AA;  
 SQ  
 Query Match 29.4%; Score 5; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 13 LRRER 17  
 Db | | | |  
 5 lrrer 9  
 RESULT 5  
 AAY42191  
 ID AAY42191 standard; peptide; 14 AA.  
 XX AC AAY42191;  
 XX 16-DEC-1999 (first entry)  
 XX Oestrogen receptor-beta MBP peptide M-70.  
 DE Oestrogen receptor beta; ER-beta; MBP; binding; detection;  
 KW ligand-dependent nuclear receptor; coactivator interaction;  
 KW hybrid system; corepressor; ligand binding domain, LBD.  
 XX Synthetic.  
 OS

XX WO9950664-A1.  
 XX 07-OCT-1999.  
 XX 01-APR-1999; 99WO-US07168.  
 XX 01-APR-1998; 98US-0053611.  
 XX (GLAX ) GLAXO GROUP LTD.  
 XX Northrop JP, Hart CP, Schatz PJ;  
 XX WPI; 1999-591374/50.  
 XX Identifying agonists or antagonists for ligand-induced activation of  
 PT predetermined nuclear receptor having ligand binding domain (LBD)  
 XX Claim 7; Fig 16; 67pp; English.  
 XX The present invention describes the use of a ligand binding domain (LBD)  
 CC transcription factor (TRX) fusion protein in methods for identifying  
 CC agonists or antagonists for ligand-induced activation of a predetermined  
 CC nuclear receptor having a LBD. Also describes are: (1) a positive hybrid  
 CC nuclear receptor signal transduction system, comprising an intact  
 CC eukaryotic host cell; (2) a nuclear receptor signal transduction system,  
 CC comprising a reverse hybrid reporter host cell; (3) a multiplexed format  
 CC assay for identifying nuclear receptor ligands; (4) identifying a  
 CC candidate pharmaceutical agent from a library of test agents; (5) a  
 CC polypeptide comprising a binding amino acid sequence that is  
 CC non-naturally occurring in a nuclear protein, and predetermined to bind  
 CC to a nuclear receptor, typically by interaction with a binding  
 CC interface of a nuclear receptor at which naturally occurring  
 CC coactivators and/or corepressors bind; and (6) an agonist ligand of a  
 CC nuclear receptor identified using above methods. The new methods are  
 CC useful for identifying agonists or antagonists for ligand-induced  
 CC activation of a predetermined nuclear receptor having a LBD. AAY42186  
 CC to AAY42204 represent specifically claimed examples of (5).  
 XX Sequence 14 AA;  
 SQ  
 Query Match 29.4%; Score 5; DB 20; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 SLLHS 8  
 Db | | | |  
 6 silhs 10  
 RESULT 6  
 AAR30448  
 ID AAR30448 standard; peptide; 16 AA.  
 XX AC AAR30448;  
 XX 06-MAY-1993 (first entry)  
 XX C242:11 MAb kappa chain CDRL.  
 DE Kappa; chain; heavy; complementarity determining region; CDR; MAb;  
 KW monoclonal antibody; C242:11; murine; IgG; hybridoma; cell line;  
 KW spleen; human; colonic; adenocarcinoma; myeloma; Sp2/0; antigen;  
 KW endocytosis.  
 XX Synthetic.  
 XX OS  
 XX EP521842-A.  
 XX 07-JAN-1993.  
 XX 03-JUL-1992; 92EP-0850166.

XX 03-JUL-1991; 91SE-0002074.  
 PR (KABI ) KABI PHARMACIA AB.  
 XX Holmgren J, Lind P, Lindholm L;  
 PI WPI; 1993-002345/01.  
 XX Monoclonal antibody reacting with CA-242 antigen - obtd. by  
 PT culturing hybridoma cell line C242:11 or mutants, useful for  
 PT diagnosis and therapy of pancreatic or colorectal cancers  
 XX Claim 5; Page 11; 15pp; English.  
 PS The sequences given in AAR30448-50 represent the kappa chain and  
 CC AAR30451-53 the heavy chain complementarity determining regions (CDR)  
 CC of a monoclonal antibody (MAB), C242:11. C242:11 is a monoclonal  
 CC murine Ab of IgG class produced when culturing in an appropriate  
 CC medium a hybridoma cell line obtained by fusing spleen cells from a  
 CC mouse, which has been immunised with a human colonic adenocarcinoma  
 CC cell line, with the murine myeloma cell line Sp2/0. C242:11 when  
 CC bound to a cell surface antigen is capable of being endocytosed or  
 CC internalised into cells.  
 XX Sequence 16 AA;  
 SQ

Query Match 29.4%; Score 5; DB 14; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 SLLHS 8  
 DB 5 silhs 9

RESULT 7  
 AAY14404  
 ID AAY14404 standard; peptide; 16 AA.  
 XX AAY14404;  
 AC  
 XX 17-AUG-1999 (first entry)  
 DT  
 DE Peptide CDR-L1 derived from anti-HCV Ser/Thr protease MAB 8D4.  
 XX Complementarity determining region; CDR; monoclonal antibody; MAB;  
 KW Hepatitis C virus; HCV; protease; binding site.  
 XX Synthetic.  
 OS  
 XX JP11127861-A.  
 PN  
 XX 18-MAY-1999.  
 PD  
 XX 29-OCT-1997; 97JP-0297451.  
 PF  
 XX 29-OCT-1997; 97JP-0297451.  
 PR  
 XX (NIHA ) JAPAN ENERGY CORP.  
 PA  
 XX WPI; 1999-350322/30.  
 DR  
 XX Neutralized antibody partial peptide derived from hepatitis C virus  
 PT - useful for inhibiting Hepatitis C Virus (HCV) serine protease  
 PT activity  
 XX Example 1; Page 13; 32pp; Japanese.  
 PS This sequence represents a peptide derived for the sequence of the light  
 XX chain variable region complementarity determining region (CDR)-1 of the  
 CC anti-hepatitis C virus (HCV) Ser/Thr protease monoclonal antibody (MAB)

CC 8D4 protein. The invention relates to the use of partial peptides  
 CC (AAY14348-Y14353) from the MAB 8D4 for inhibiting HCV serine protease  
 CC activity.  
 XX Sequence 16 AA;  
 SQ

Query Match 29.4%; Score 5; DB 20; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 SLLHS 8  
 DB 5 silhs 9

RESULT 8  
 AAY95222  
 ID AAY95222 standard; Peptide; 16 AA.  
 XX AAY95222;  
 AC AAY95222;  
 XX 29-AUG-2000 (first entry)  
 DT  
 DE Anti-platelet glycoprotein Ib human H1b-5 VL CDR1.  
 XX  
 XX Variable light chain; single chain antibody; scFv; human; H1b-5;  
 KW glycoprotein Ib alpha; platelet; aggregation; antiaggregant;  
 KW antithrombotic; thrombus; therapy; diagnostic; CDR1;  
 KW complementarity determining region.  
 XX Homo sapiens.  
 OS  
 XX WO200026667-A1.  
 PN  
 XX 11-MAY-2000.  
 PD  
 XX 29-OCT-1999; 99WO-US25495.  
 PF  
 XX 30-OCT-1998; 98US-0106275.  
 PR  
 XX (MILL/) MILLER J L.  
 PA  
 PI Miller JL;  
 XX  
 XX WPI; 2000-365744/31.  
 DR  
 XX Isolated nucleic acid molecule encoding anti-human platelet  
 PT glycoprotein Ib alpha molecule useful for producing antibodies which  
 PT inhibit platelet aggregation -  
 XX Claim 19; Fig 8; 89pp; English.  
 PS The present sequence is that of complementarity determining region  
 CC 1 (CDR1) of the light chain variable region (VL) of human single  
 CC chain antibody (scFv) H1b-5, which is directed against platelet  
 CC glycoprotein Ib (GP1b). The H1b series of scFv antibodies was  
 CC isolated from a human synthetic VH and VL scFv library on the  
 CC basis of their binding to platelet GP1b. Whether displayed as  
 CC surface proteins on a phagemid or secreted as free scFv by  
 CC Escherichia coli, the H1b scFv clones are capable of inhibiting  
 CC von Willebrand factor-dependent aggregation of platelets. The scFv  
 CC are composed of native human protein sequences and are therefore  
 CC attractive potential reagents for therapeutic purposes. They  
 CC provide a new class of antithrombotic agents, useful for the  
 CC prevention of platelet-dependent thrombi in diseased arteries,  
 CC bypass grafts, dialysis etc., and can also be used as diagnostic  
 CC reagents. Methods of inhibiting aggregation of platelets, of  
 CC binding human platelet GP1b alpha and of selecting a VH or VL  
 CC region of an antibody that inhibits platelet aggregation are  
 CC claimed. Fragments of the scFv VH or VL chain, including CDR  
 CC fragments, are also claimed.  
 XX

SQ Sequence 16 AA;

Query Match 29.4%; Score 5; DB 21; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 SLLHS 8  
 |||||  
 Db 5 silhs 9

RESULT 9  
 ID AAW83469 standard; peptide; 4 AA.  
 XX AAW83469;  
 AC  
 XX  
 DT 08-MAR-1999 (first entry)  
 XX  
 DE Human growth hormone mutant peptide #29.  
 XX  
 KW Human; growth hormone; hGH; phagemid particle; enzyme substrate;  
 KW fusion gene; phage protein coat.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN US5846765-A.  
 XX  
 XX 08-DEC-1998.  
 PD  
 XX  
 PF 16-MAY-1995; 95US-0441871.  
 XX  
 XX 03-DEC-1993; 93US-0161692.  
 PR 03-DEC-1990; 90US-0621667.  
 PR 10-APR-1991; 91US-0683400.  
 PR 14-JUN-1991; 91US-0715300.  
 PR 09-AUG-1991; 91US-0743614.  
 PR 03-DEC-1991; 91WO-US091133.  
 PR 06-APR-1992; 92US-0864452.  
 PR 30-APR-1993; 93US-0050058.  
 PR 05-APR-1995; 93US-0418928.  
 PR 16-MAY-1995; 95US-0441871.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Matthews DJ, Wells JA, Zoller MJ;  
 XX  
 DR WPI; 1999-059058/05.  
 XX  
 PT Selection of polypeptide substrates - using phagemid particles  
 PT displaying poly-peptide(s) as coat protein fusions  
 XX  
 XX Example 8; Column 32; 81pp; English.  
 XX  
 CC The present invention describes a method for selecting novel polypeptide  
 CC substrates. The method comprises: (a) constructing a replicable  
 CC expression vector comprising a transcription regulatory element operably  
 CC linked to a gene fusion; (b) mutating the vector at one or more selected  
 CC positions within the second gene thereby forming a family of related  
 CC plasmids encoding substrate peptides; (c) transforming suitable host  
 CC cells with the plasmids; (d) infecting the transformed host cells with a  
 CC helper phage having a gene encoding the phage coat protein; (e) culturing  
 CC the transformed infected host cells under conditions suitable for forming  
 CC recombinant phagemid particles containing at least a portion of the  
 CC plasmid and capable of transforming the host, the conditions being  
 CC adjusted so that no more than a minor amount of phagemid particles  
 CC display more than one copy of the fusion protein on the surface of the  
 CC particle; (f) exposing the phagemid particles to a process capable of  
 CC modifying at least one covalent bond of an amino acid in the substrate  
 CC peptide of at least a portion of the phagemid particles; (g) contacting  
 CC the family of exposed phagemid particles with an affinity molecule, where

CC the affinity molecule has affinity for the amino acid residue having the  
 CC modified covalent bond; and (h) separating the phagemid particles that  
 CC bind to the affinity molecule from those that do not bind. The selection  
 CC method is used for identifying enzyme substrates. The present sequence  
 CC represents a human growth hormone mutant peptide from an example of the  
 CC present invention.

XX  
 SQ Sequence 4 AA;

Query Match 23.5%; Score 4; DB 20; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 SSKE 11  
 |||||  
 Db 1 sske 4

RESULT 10  
 AAR62114  
 ID AAR62114 standard; peptide; 5 AA.  
 XX  
 AC AAR62114;  
 XX  
 DT 27-APR-1995 (first entry)  
 XX  
 DE Hydrophilic motif from U1 snRNP 70K protein.  
 XX  
 KW Small ribonucleoprotein complex; U1 snRNP; 70K protein; epitope;  
 KW autoantibody; immunoinfective cluster virus; nuclear protein antigen;  
 KW systemic rheumatic disorder; human immunodeficiency virus; HIV-1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9420141-A.  
 XX  
 PD 15-SEP-1994.  
 XX  
 PF 10-MAR-1994; 94WO-US02631.  
 XX  
 PR 11-MAR-1993; 93US-0029850.  
 XX  
 XX (UYSC-) UNIV SOUTHERN CALIFORNIA.  
 XX  
 PI Douvas A, Ehresmann G, Takehana Y;  
 XX  
 DR WPI; 1994-302689/37.  
 XX  
 PT Methods for treating immunoinfective cluster virus infections -  
 PT utilise antibodies or fragments characteristic of auto antibodies  
 PT produced by patients with rheumatic disorders  
 XX  
 XX Disclosure; Page 8; 106pp; English.  
 XX  
 CC The sequence RERRR (AAR62113) is a preferred example of an alternating  
 CC acidic/basic amino acid, hydrophilic epitope motif, found in the  
 CC U1 snRNP 70K protein. It also occurs as RREER and EREER (AAR62114  
 CC and AAR62115) in the 70K protein. The motif is also found in similar  
 CC form in immunoinfective cluster viruses. The motif serves as an  
 CC epitope for anti-viral antibodies and also for autoantibodies which  
 CC occur in high titre in patients suffering from systemic rheumatic  
 CC disorders. Sera from such patients could be used for treatment of  
 CC immunoinfective cluster virus (e.g. HIV,  
 CC EBV, rubella virus) infections.  
 XX  
 SQ Sequence 5 AA;

Query Match 23.5%; Score 4; DB 15; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 RRER 17  
 Db 1 rrer 4

## RESULT 11

AAW00407  
 ID AAW00407 standard; peptide; 5 AA.

XX AC AAW00407;

XX DT 29-AUG-1996 (first entry)

XX DE Interleukin-6 antagonist peptide.

XX KW IL-6; antagonist; autoimmune disease.

XX OS Synthetic.

XX PN JP07324097-A.

XX PD 12-DEC-1995.

XX PF 30-MAY-1994; 94JP-0117259.

XX PR 30-MAY-1994; 94JP-0117259.

XX PA (DAIL ) DAICEL CHEM IND LTD.

XX PA (FUJI ) FUJISAWA PHARM CO LTD.

XX DR WPI; 1996-065476/07.

XX PT Interleukin 6 antagonist - useful for treating autoimmune diseases

XX PS Claims 2, 6; Pages 2, 3; 19pp; Japanese.

XX CC New IL-6 antagonists are provided which are of formula X-W-Y, in which X is H or an amino-protecting group, Y is OH or a carboxy-protecting group, and W is a peptide containing all or part of the sequence as given in AAW00401, AAW00402, AAW00403 or AAW00404, where any free mercapto groups in the sequence are optionally protected. The present sequence is a specifically preferred partial sequence of AAW00401 and is itself claimed as a new chemical entity.

XX CC The IL-6 antagonists are useful for treating autoimmune diseases.

XX SQ Sequence 5 AA;

Query Match 23.5%; Score 4; DB 17; Length 5;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SSKE 11  
 Db 2 sske 5

## RESULT 12

AAW61650  
 ID AAW61650 standard; peptide; 5 AA.

XX AC AAW61650;

XX DT 27-OCT-1998 (first entry)

XX DE Human reflex tears protein (cystatin S).

XX KW Human; non-ocular disease; tear; cancer; breast; prostate.

XX OS Homo sapiens.

XX PN WO9835229-A1.

XX XX

PD 13-AUG-1998.

XX PF 06-FEB-1998; 98WO-AU000071.

XX PR 07-FEB-1997; 97AU-0005009.

XX PA (MACQ-) MACQUARIE RES LTD.

XX PA (UNIX ) UNISEARCH LTD.

XX PI Bolis S, Gooley AA, Herbert B, Molloy M, Morris C;

XX PI Walsh B, Willcox M, Williams KL;

XX DR WPI; 1998-447373/38.

XX PT Screening for non-ocular disease - by analysing tears for marker proteins, particularly indicative of cancer and genetic disease, also new proteins and nucleic acid encoding them

XX PS Disclosure; Page 7; 14pp; English.

XX CC The peptides AAW61650-W61652 were identified in human reflex tears, in the method of the invention for screening for, or detecting, non-ocular disease by analysing tears. Biochemicals, specifically proteins, are isolated from tears, particularly by chromatography or electrophoresis, especially two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), then detected, e.g. with labelled specific reagents, in (radio)immunoassay. The method is used to detect cancer, particularly of breast or prostate, or a genetic disease, in humans or animals.

XX SQ Sequence 5 AA;

Query Match 23.5%; Score 4; DB 19; Length 5;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SSKE 11  
 Db 2 sske 5

## RESULT 13

AAW62104

ID AAR62104 standard; peptide; 6 AA.

XX AC AAR62104;

XX DT 27-APR-1995 (first entry)

XX DE Hydrophilic motif from nuclear protein antigens.

XX KW Small ribonucleoprotein complex; U1 snRNP; 70K protein; epitope; autoantibody; immunoinfective cluster virus; nuclear protein antigen; systemic rheumatic disorder; human immunodeficiency virus; HIV-1; centromere CENP-B; thyroglobulin-h; thyroid peroxidase; scleroderma; systemic lupus erythematosus.

XX OS Homo sapiens.

XX PN WO9420141-A.

XX PD 15-SEP-1994.

XX PF 10-MAR-1994; 94WO-US02631.

XX PR 11-MAR-1993; 93US-0029850.

XX PA (UYSC-) UNIV SOUTHERN CALIFORNIA.

XX PI Douvas A, Ehresmann G, Takehana Y;

XX DR WPI; 1994-302689/37.

Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KEKL 7  
Db 5 KEKL 8

Search completed: September 15, 2001, 12:57:49  
Job time: 108 sec

RESULT 13  
A45199  
L-hyosiphorin - Java medaka (fragment)  
C:Species: Oryzias latipes (Java medaka)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-May-1999  
C:Accession: A45199  
R:Taguchi, T.; Seko, A.; Kitajima, K.; Inoue, S.; Iwamatsu, T.; Khoo, K.H.; Morris, H.R.  
J. Biol. Chem. 268, 2353-2362, 1993  
A:Title: Structural studies of a novel type of tetraantennary sialoglycan unit in a carb  
A:Reference number: A45199; MUID:93155036  
A:Accession: A45199  
A:Molecule type: protein  
A:Residues: 1-9 <TAG>  
C:Comment: This sialoglycoprotein is 90% carbohydrate by weight. This low molecular weigh  
C:Keywords: glycoprotein  
F:5/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 8.0%; Score 4; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 DAAS 35  
Db 1 DAAS 4

RESULT 14  
PS0252  
16K protein 5404 - rice (strain Nihonbare) (fragment)  
C:Species: Oryza sativa (rice)  
C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 05-Apr-1995  
C:Accession: PS0252  
R:Tugita, A.  
submitted to JIPID, April 1993  
A:Reference number: PS0206  
A:Accession: PS0252  
A:Molecule type: protein  
A:Residues: 1-14 <TSU>  
A:Experimental source: strain Nihonbare  
C:Comment: Molecular weight 16K, pI 4.9.

Query Match 8.0%; Score 4; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 VLEA 39  
Db 2 VLEA 5

RESULT 15  
PC4382  
dehydrin 4.5K polypeptide - Soybean (fragment)  
N:Alternate names: acid soluble 26K protein  
C:Species: Glycine max cv. Nattosoryu (soybean)  
C:Date: 06-Nov-1997 #sequence\_revision 06-Nov-1997 #text\_change 07-May-1999  
C:Accession: PC4382  
R:Monma, M.; Haraguchi, K.; Saito, M.; Chikuni, K.; Harada, K.  
Biosci. Biotechnol. Biochem. 61, 1286-1291, 1997  
A:Title: Purification and characterization of the acid soluble 26-kDa polypeptide from s  
A:Reference number: PC4380; MUID:97446521  
A:Accession: PC4382  
A:Molecule type: protein  
A:Residues: 1-14 <WOM>  
A:Experimental source: seed  
C:Comment: This protein is rich in hydrophilic amino acid, and highly heat stable.

Query Match 8.0%; Score 4; DB 2; Length 14;

T07311  
hypothetical protein 47b - Chlorella vulgaris chloroplast  
C:Species: chloroplast Chlorella vulgaris  
A:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000  
C:Accession: T07311  
R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakas  
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlo  
A:Reference number: 215985; MUID:97303241  
A:Accession: T07311  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-47 <NA>  
A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57959.1; PID:g2224475  
C:Genetics:  
A:Genome: chloroplast  
C:Keywords: chloroplast

Query Match 10.0%; Score 5; DB 2; Length 47;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 SSKEK 6  
Db 3 SSKEK 7

RESULT 9  
S61469  
p83/100 protein - Borrelia afzelii (strain PLe and others) (fragment)  
C:Species: Borrelia afzelii  
A:Variety: strain PLe and others  
C:Date: 19-Mar-1997 #sequence\_revision 13-Mar-1998 #text\_change 08-Oct-1999  
R:Roessler, D.; Eifert, H.; Jauris-Helpke, S.; Lehnert, G.; Preac-Mursic, V.; Teepe, J.  
Med. Microbiol. Immunol. 184, 23-32, 1995  
A:Title: Molecular and immunological characterization of the p83/100 protein of various  
A:Reference number: S61461; MUID:96149106  
A:Accession: S61469  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-48 <ROE>  
A:Cross-references: EMBL:X81529; NID:g928973; PIDN:CAA57248.1; PID:g928974; EMBL:X81530;  
A:Experimental source: strain PLe; strain PKJ7; strain PGau  
C:Keywords: surface antigen

Query Match 10.0%; Score 5; DB 2; Length 48;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 SSKEK 6  
Db 27 SSKEK 31

RESULT 10  
S61472  
p83/100 protein - Borrelia afzelii (strain PWudi) (fragment)  
C:Species: Borrelia afzelii  
A:Variety: strain PWudi  
C:Date: 19-Mar-1997 #sequence\_revision 05-Dec-1998 #text\_change 08-Oct-1999  
R:Roessler, D.; Eifert, H.; Jauris-Helpke, S.; Lehnert, G.; Preac-Mursic, V.; Teepe, J.  
Med. Microbiol. Immunol. 184, 23-32, 1995  
A:Title: Molecular and immunological characterization of the p83/100 protein of various  
A:Reference number: S61461; MUID:96149106  
A:Accession: S61472  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-48 <ROE>  
A:Cross-references: EMBL:X81535

Query Match 10.0%; Score 5; DB 2; Length 48;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 SSKEK 6  
Db 27 SSKEK 31

RESULT 11  
B33932  
Ig mu chain D region (D23) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 16-Aug-1996  
C:Accession: B33932  
R:Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.  
Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989  
A:Title: Two murine natural polyclonal reactive autoantibodies are encoded by nonmutated ger  
A:Reference number: A33932; MUID:89282823  
A:Accession: B33932  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-6 <BAC>  
A:Cross-references: GB:M27107  
C:Keywords: immunoglobulin

Query Match 8.0%; Score 4; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 EKLR 8  
Db 1 EKLR 4

RESULT 12  
PT0560  
T-cell receptor beta chain V-D-J region (126-1CE) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0560  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions  
A:Reference number: PT0509; MUID:91277601  
A:Accession: PT0560  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-6 <FE>  
A:Experimental source: day 18 fetal thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 8.0%; Score 4; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 34 ASVL 37  
Db 1 ASVL 4

A:Experimental source: strain PWudi  
R:Roessler, D.  
submitted to the EMBL Data Library, September 1994  
A:Reference number: S72307  
A:Accession: S72307  
A:Molecule type: DNA  
A:Residues: 1-7, 'F', 8-48 <ROW>  
A:Cross-references: EMBL:X81535; NID:g928977; PIDN:CAA57254.1; PID:g928978  
A:Experimental source: strain PWudi  
C:Keywords: surface antigen

Query Match 10.0%; Score 5; DB 2; Length 48;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 SSKEK 6  
Db 27 SSKEK 31

RESULT 11  
B33932  
Ig mu chain D region (D23) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 16-Aug-1996  
C:Accession: B33932  
R:Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.  
Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989  
A:Title: Two murine natural polyclonal reactive autoantibodies are encoded by nonmutated ger  
A:Reference number: A33932; MUID:89282823  
A:Accession: B33932  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-6 <BAC>  
A:Cross-references: GB:M27107  
C:Keywords: immunoglobulin

Query Match 8.0%; Score 4; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EKLR 8  
Db 1 EKLR 4

RESULT 12  
PT0560  
T-cell receptor beta chain V-D-J region (126-1CE) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0560  
R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions  
A:Reference number: PT0509; MUID:91277601  
A:Accession: PT0560  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-6 <FE>  
A:Experimental source: day 18 fetal thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 8.0%; Score 4; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ASVL 37  
Db 1 ASVL 4



Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
Best Local Similarity 100.0%; DB 2; Length 31;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 16 LRTLL 20

RESULT 3  
I61697  
myosin - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 05-Nov-1999  
C:Accession: I61697  
R:Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.  
Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994  
A:Title: Identification and overlapping expression of multiple unconventional myosin genes  
A:Reference number: A55758; MUID:94294418  
A:Accession: I61697  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-31 <RES>  
A:Cross-references: GB:L29146; NID:g457254; PIDN:AAA20909.1; PID:g531139

Query Match 10.0%; Score 5; DB 2; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VLEAT 40  
Best Local Similarity 100.0%; DB 2; Length 28;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 24 VLEAT 28

RESULT 4  
I46598  
myosin - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 05-Nov-1999  
C:Accession: I46598  
R:Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.  
Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994  
A:Title: Identification and overlapping expression of multiple unconventional myosin genes  
A:Reference number: A55758; MUID:94294418  
A:Accession: I46598  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-31 <BEM>  
A:Cross-references: GB:L29133; NID:g457343; PIDN:AAA20918.1; PID:g531149

Query Match 10.0%; Score 5; DB 2; Length 31;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VLEAT 40  
Best Local Similarity 100.0%; DB 2; Length 28;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 24 VLEAT 28

RESULT 5  
S27307  
surface-array protein - Aeromonas hydrophila (fragment)  
C:Species: Aeromonas hydrophila  
C>Date: 22-Nov-1993 #sequence\_revision 09-Mar-1996 #text\_change 07-May-1999  
C:Accession: S27307; A33184  
R:Kokka, R.P.; Vedros, N.A.; Janda, J.M.  
J. Gen. Microbiol. 138, 1229-1236, 1992  
A:Title: Immunochemical analysis and possible biological role of an Aeromonas hydrophila  
A:Reference number: S27307; MUID:94207495  
A:Accession: S27307  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-35 <K02>

Query Match 10.0%; Score 5; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 DAASV 36  
Best Local Similarity 100.0%; DB 2; Length 35;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 DAASV 9

RESULT 6  
A82382  
hypothetical protein VCA1064 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: A82382  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.;  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers  
I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833  
A:Accession: A82382  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-40 <HEI>  
A:Cross-references: GB:AE004432; GB:AE003853; NID:g9658509; PIDN:AAF96958.1; GSPDB:GN  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VCA1064  
A:Map position: 2

Query Match 10.0%; Score 5; DB 2; Length 40;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KGRKN 31  
Best Local Similarity 100.0%; DB 2; Length 40;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 35 KGRKN 39

RESULT 7  
T07248  
hypothetical protein 42a - Chlorella vulgaris chloroplast  
C:Species: chloroplast Chlorella vulgaris  
C>Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000  
C:Accession: T07248  
R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Na  
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga C  
A:Reference number: Z15985; MUID:97303241  
A:Accession: T07248  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-42 <WAK>  
A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57895.1; PID:g2224411  
C:Genetics:  
A:Genome: chloroplast  
C:Keywords: chloroplast

Query Match 10.0%; Score 5; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SKEKL 7  
Best Local Similarity 100.0%; DB 2; Length 42;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 SKEKL 23

RESULT 8

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run On: September 15, 2001, 12:56:01 ; Search time 12.53 Seconds  
(without alignments)  
303.969 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_189

Perfect score: 50

Sequence: 1 HSSKEKLRRRIKYCCQLR.....NDAASVLEATVDYVKYIREK 50

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 11081

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database :

PIR\_68:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	10.0	29	S10050	ribosomal protein
2	5	10.0	31	G82071	hypothetical prote
3	5	10.0	31	I61697	myosin - human (fr
4	5	10.0	31	I46598	myosin - pig (frag
5	5	10.0	35	S27307	surface-array prot
6	5	10.0	40	A82382	hypothetical prote
7	5	10.0	42	T07248	hypothetical prote
8	5	10.0	47	T07311	hypothetical prote
9	5	10.0	48	S61469	p83/100 protein -
10	5	10.0	48	S61472	p83/100 protein -
11	4	8.0	6	B33932	Ig mu chain D regi
12	4	8.0	6	PT0560	T-cell receptor be
13	4	8.0	9	A45199	L-hyosophorin - Ja
14	4	8.0	14	P50252	16K protein 5404 -
15	4	8.0	14	PC4382	dehydrin 4.5K poly
16	4	8.0	15	B49177	21K protein p2, mi
17	4	8.0	17	A05168	conantokin G [vali
18	4	8.0	17	E33113	gingipain, 44K - P
19	4	8.0	18	QHEC2	heat-stable entero
20	4	8.0	18	S74195	epoxide hydrolase
21	4	8.0	18	A60103	heat-stable entero
22	4	8.0	19	B61056	cytochrome P450 PB
23	4	8.0	20	I38417	HLA-A11 - human (f
24	4	8.0	20	S10680	probable 7-ethoxyc
25	4	8.0	20	C49753	hypothetical prote
26	4	8.0	20	S18582	hypothetical prote
27	4	8.0	20	T50757	pufk protein (limp
28	4	8.0	20	A60530	hyosophorin - Japa
29	4	8.0	22	PT0431	leucyl aminopeptid

30 4 8.0 22 2 S00189  
31 4 8.0 22 2 B33174  
32 4 8.0 22 2 H49410  
33 4 8.0 23 2 I48936  
34 4 8.0 23 2 I49413  
35 4 8.0 23 2 S31210  
36 4 8.0 23 2 PH1641  
37 4 8.0 24 2 A84023  
38 4 8.0 25 2 S56002  
39 4 8.0 27 2 A35948  
40 4 8.0 27 2 I47719  
41 4 8.0 27 2 PC4234  
42 4 8.0 27 2 B38746  
43 4 8.0 27 2 C38746  
44 4 8.0 27 2 C38746  
45 4 8.0 27 2 F38746

## ALIGNMENTS

## RESULT 1

S10050

ribosomal protein L19.e - fission yeast (Schizosaccharomyces pombe) (fragment)  
N:Alternate names: ribosomal protein SP-L15  
C:Species: Schizosaccharomyces pombe

C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 30-Sep-1993  
C:Accession: S10050

R:Otaka, E.; Higo, K.I.; Itoh, T.

Mol. Gen. Genet. 191, 519-524, 1983

A:Title: Yeast ribosomal proteins: VII. Cytoplasmic ribosomal proteins from Schizosac  
A:Reference number: S07293; MUID:84038947

A:Accession: S10050

A:Molecule type: protein

A:Residues: 1-29 &lt;OTA&gt;

C:Superfamily: rat ribosomal protein L19

C:Keywords: protein biosynthesis; ribosome

Query Match 10.0%; Score 5; DB 2; Length 29;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 AASVL 37

I I I I I

Db 10 AASVL 14

## RESULT 2

G82071

hypothetical protein VC2477 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: G82071

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.  
I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: G82071

A&gt;Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-31 &lt;HEI&gt;

A:Cross-references: GB:AE004317; GB:AE003852; NID:g9657050; PIDN:AAF95619.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2477

A:Map position: 1

Query Match

Best Local Similarity 100.0%; Score 5; DB 2; Length 31;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**THIS PAGE BLANK (USPTO)**



QY 19 LRTLL 23  
|||||  
Db 4 LRTLL 8

## RESULT 13

US-08-433-613-9  
; Sequence 9, Application US/08433613A  
; Patent No. 6162434

## GENERAL INFORMATION:

; APPLICANT: Buelow, Roland  
; TITLE OF INVENTION: Cytomodulating Peptide for Inhibiting Lymphocyte  
; FILE REFERENCE: A61008/RET/TAL  
; CURRENT APPLICATION NUMBER: US/08/433,613A  
; CURRENT FILING DATE: 1995-05-03  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-08-433-613-9

Query Match 10.0%; Score 5; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
|||||  
Db 8 LRTLL 12

## RESULT 14

US-08-433-613-48  
; Sequence 48, Application US/08433613A  
; Patent No. 6162434

## GENERAL INFORMATION:

; APPLICANT: Buelow, Roland  
; TITLE OF INVENTION: Cytomodulating Peptide for Inhibiting Lymphocyte  
; FILE REFERENCE: A61008/RET/TAL  
; CURRENT APPLICATION NUMBER: US/08/433,613A  
; CURRENT FILING DATE: 1995-05-03  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-08-433-613-48

Query Match 10.0%; Score 5; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
|||||  
Db 8 LRTLL 12

## RESULT 15

US-08-750-856A-8  
; Sequence 8, Application US/08750856A  
; Patent No. 5858672

## GENERAL INFORMATION:

; APPLICANT: SONIGO, PIERRE  
; APPLICANT: PANCINO, GIANFRANCO  
; APPLICANT: PETERHANS, ERNST  
; APPLICANT: BERTONI, GIUSEPPE  
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS AND CORRESPONDING  
; TITLE OF INVENTION: PEPTIDE FRAGMENTS FROM THE CAPRINE ARTHRITIS-ENCEPHALITIS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/750,856A  
; FILING DATE: 30-DEC-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 94/07933  
; FILING DATE: 28-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 917-052-0 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-750-856A-8

Query Match 10.0%; Score 5; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VLEAT 40  
|||||  
Db 2 VLEAT 6

Search completed: September 15, 2001, 12:57:02  
Job time: 106 sec

APPLICANT: Buelow, Roland  
TITLE OF INVENTION: Treatment for Inhibiting the Progression  
of Autoimmune Disease  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,504A  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-60130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 949-8711  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-440-504A-9

Query Match 10.0%; Score 5; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
Db 4 LRTLL 8

RESULT 11  
US-08-440-504A-14  
Sequence 14, Application US/08440504A  
Patent No. 5753625  
GENERAL INFORMATION:  
APPLICANT: Buelow, Roland  
TITLE OF INVENTION: Treatment for Inhibiting the Progression  
of Autoimmune Disease  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,504A  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.

REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-60130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 949-8711  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-440-504A-14

Query Match 10.0%; Score 5; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
Db 4 LRTLL 8

RESULT 12  
PCT-US94-12985-13  
Sequence 13, Application PC/TUS9412985  
GENERAL INFORMATION:  
APPLICANT: The Board of Trustees for the Leland Stanford Junior  
University  
TITLE OF INVENTION: SURFACE MEMBRANE PROTEINS AND THEIR  
EFFECT ON IMMUNE RESPONSE  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/12985  
FILING DATE: 10-NOV-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/150,493  
FILING DATE: 10-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Rowland, Bertram I  
REGISTRATION NUMBER: 20,015  
REFERENCE/DOCKET NUMBER: FP-58976-PC/BIR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US94-12985-13

Query Match 10.0%; Score 5; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/105,416  
;; FILING DATE: 12-AUG-1993  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Rowland, Bertram I.  
;; REGISTRATION NUMBER: 20015  
;; REFERENCE/DOCKET NUMBER: A55115-4/BIR  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-781-1989  
;; TELEFAX: 415-398-3249  
;; INFORMATION FOR SEQ ID NO: 36:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 7 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-105-416-36

Query Match 10.0%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 19 LRTLL 23  
Db 1 LRTLL 5

RESULT 8  
US-08-473-656A-36  
; Sequence 36, Application US/08473656A  
; Patent No. 5853999  
; GENERAL INFORMATION:  
; APPLICANT: Olsson, Lennart  
; APPLICANT: Goodenow, Robert S  
; APPLICANT: Goldstein, Avram  
; TITLE OF INVENTION: Class I MHC Modulation of Surface  
; TITLE OF INVENTION: Receptor Activity  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fierh, Hohbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/473,656A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/105,416  
; FILING DATE: 12-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rowland, Bertram I.  
; REGISTRATION NUMBER: 20015  
; REFERENCE/DOCKET NUMBER: A55115-4/BIR  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid

;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-473-656A-36  
Query Match 10.0%; Score 5; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 19 LRTLL 23  
Db 1 LRTLL 5  
RESULT 9  
US-08-222-851-14  
; Sequence 14, Application US/08222851  
; Patent No. 5723128  
; GENERAL INFORMATION:  
; APPLICANT: CLAYBERGER, CAROL A.  
; APPLICANT: KRENSKY, ALAN M.  
; APPLICANT: PARHAM, PETER  
; TITLE OF INVENTION: CYTOTOXIC T-CELL LYMPHOCYTE ("CTL")  
; TITLE OF INVENTION: ACTIVITY REGULATION BY CLASS I MHC PEPTIDES  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 PENNSYLVANIA AVENUE, NW, STE 5500  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/222,851  
; FILING DATE: 05-APR-1994  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MILLMAN, ROBERT A.  
; REGISTRATION NUMBER: 36,217  
; REFERENCE/DOCKET NUMBER: 28600-20200.22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 494-0792  
; TELEX: 90-4030 MRSNFOERSWSH  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-222-851-14

Query Match 10.0%; Score 5; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 19 LRTLL 23  
Db 4 LRTLL 8  
RESULT 10  
US-08-440-504A-9  
; Sequence 9, Application US/08440504A  
; Patent No. 5753625  
; GENERAL INFORMATION:

;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;; ORGANISM: Salmonella typhimurium  
US-07-944-143C-17

Query Match 10.0%; Score 5; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LRTLL 23  
|||||  
Db 2 LRTLL 6

RESULT 5  
PCT-US93-08214-17  
; Sequence 17, Application PC/TUS9308214  
; GENERAL INFORMATION:  
; APPLICANT: Oklahoma Medical Research Foundation  
; TITLE OF INVENTION: Peptide Diagnostics and Therapeutics for Spondyloarthropat  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street, Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: United States  
; ZIP: 30309-4530

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08214  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Padst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 815-6508  
TELEFAX: (404) 815-6555  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Salmonella typhimurium

Query Match 10.0%; Score 5; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LRTLL 23  
|||||  
Db 2 LRTLL 6

RESULT 6  
PCT-US93-12679-5  
; Sequence 5, Application PC/TUS9312679  
; GENERAL INFORMATION:  
; APPLICANT: Rath, Matthias  
; TITLE OF INVENTION: SYNTHETIC OLIGOPEPTIDES ANALOGOUS TO  
; TITLE OF INVENTION: PROTEIN SIGNAL SEQUENCES -METHODS OF IDENTIFICATION AND  
; TITLE OF INVENTION: METHODS OF USE  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHELDON & MAK  
; STREET: 401 Florence Street  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/12679  
FILING DATE: 30-DEC-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cranfill, Raymond B  
REGISTRATION NUMBER: 32,845  
REFERENCE/DOCKET NUMBER: RATH-10016PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-322-5333  
TELEFAX: 415-322-5499  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US93-12679-5

Query Match 10.0%; Score 5; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSKEK 6  
|||||  
Db 2 SSKEK 6

RESULT 7  
US-08-105-416-36  
; Sequence 36, Application US/08105416  
; Patent No. 5639558  
; GENERAL INFORMATION:  
; APPLICANT: Olsson, Lennart  
; APPLICANT: Goodenow, Robert S  
; APPLICANT: Goldstein, Avram  
; TITLE OF INVENTION: Class I MHC Modulation of Surface  
; TITLE OF INVENTION: Receptor Activity  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flerh, Hobbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible



ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/447,702  
FILING DATE: 23-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/194,588  
FILING DATE: 10-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 92-09-882  
FILING DATE: 10-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 003025-019  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-447-702-1

Query Match 12.0%; Score 6; DB 1; Length 27;  
Best Local Similarity 100.0%; Pred. No. 6.1;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 LEATVD 42  
Db 20 LEATVD 25

RESULT 3  
US-08-465-615-1  
Sequence 1, Application US/08465615  
Patent No. 5635391  
GENERAL INFORMATION:  
APPLICANT: PETRE, Dominique  
APPLICANT: CERBELEAUD, Edith  
APPLICANT: LEVY-SCHIL, Sophie  
APPLICANT: CROUZET, Joel  
TITLE OF INVENTION: POLYPEPTIDES POSSESSING A NITRILASE  
TITLE OF INVENTION: ACTIVITY, DNA SEQUENCE CODING FOR SAID POLYPEPTIDES,  
TITLE OF INVENTION: EXPRESSION CASSETTES AND HOST MICROORGANISMS ENABLING THEM  
TITLE OF INVENTION: TO BE OBTAINED, AND METHOD OF CONVERTING NITRILES TO  
TITLE OF INVENTION: CARBOXYLATES BY MEANS OF SAID POLYPEPTIDE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,615  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/194,588  
FILING DATE: 10-FEB-1994  
APPLICATION NUMBER: FR 9209882  
FILING DATE: 10-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Crane-Feury, Sharon E  
REGISTRATION NUMBER: 36,113  
REFERENCE/DOCKET NUMBER: 003025-015  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-465-615-1

Query Match 12.0%; Score 6; DB 1; Length 27;  
Best Local Similarity 100.0%; Pred. No. 6.1;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 LEATVD 42  
Db 20 LEATVD 25

RESULT 4  
US-07-944-143C-17  
Sequence 17, Application US/07944143C  
Patent No. 5719064  
GENERAL INFORMATION:  
APPLICANT: Scofield, R. Hal  
APPLICANT: Hartley, John B.  
TITLE OF INVENTION: Peptide Diagnostics and Therapeutics for  
TITLE OF INVENTION: Spondyloarthropathies  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/944,143C  
FILING DATE:  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMR138  
TELEPHONE: (404)873-8794  
TELEFAX: (404)873-8795  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:55:16 ; Search time 12.43 Seconds  
(without alignments)  
82.825 Million cell updates/sec

Title: US-09-389-000-2\_copy\_140\_189

Perfect score: 50

Sequence: 1 HSKSEKLRRERIKYCEQLR.....NDAASVLEATVDYVKVIREK 50

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 197339-seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 140039

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : Issued Patents\_AA:\*

1: /cgnl\_7/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgnl\_7/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgnl\_7/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgnl\_7/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgnl\_7/ptodata/1/iaa/PCTUS\_COMB.pep.\*

6: /cgnl\_7/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	12.0	15	4	US-08-986-837-8
2	6	12.0	27	1	US-08-447-702-1
3	6	12.0	27	1	US-08-465-615-1
4	5	10.0	6	1	US-07-944-143C-17
5	5	10.0	6	5	PCT-US93-08214-17
6	5	10.0	6	5	PCT-US93-12679-5
7	5	10.0	7	1	US-08-105-416-36
8	5	10.0	7	2	US-08-473-656A-36
9	5	10.0	10	1	US-08-222-851-14
10	5	10.0	10	1	US-08-440-504A-9
11	5	10.0	10	1	US-08-440-504A-14
12	5	10.0	10	5	PCT-US94-12985-13
13	5	10.0	14	4	US-08-433-613-9
14	5	10.0	14	4	US-08-433-613-48
15	5	10.0	15	2	US-08-750-856A-8
16	5	10.0	16	1	US-07-944-143C-16
17	5	10.0	16	5	PCT-US93-08214-16
18	5	10.0	17	1	US-07-976-872B-1
19	5	10.0	17	1	US-07-976-872B-2
20	5	10.0	17	1	US-07-976-872B-3
21	5	10.0	17	1	US-07-976-872B-4
22	5	10.0	17	1	US-08-105-416-12
23	5	10.0	17	1	US-08-105-416-13
24	5	10.0	17	1	US-08-105-416-18
25	5	10.0	17	1	US-08-105-416-19
26	5	10.0	17	1	US-08-105-416-20
27	5	10.0	17	1	US-08-105-416-27

28 5 10.0 17 1 US-08-105-416-29 Sequence 29, Appl  
29 5 10.0 17 1 US-08-105-416-30 Sequence 30, Appl  
30 5 10.0 17 1 US-08-105-416-37 Sequence 37, Appl  
31 5 10.0 17 1 US-08-105-416-38 Sequence 38, Appl  
32 5 10.0 17 2 US-08-473-656A-12 Sequence 12, Appl  
33 5 10.0 17 2 US-08-473-656A-13 Sequence 13, Appl  
34 5 10.0 17 2 US-08-473-656A-18 Sequence 18, Appl  
35 5 10.0 17 2 US-08-473-656A-19 Sequence 19, Appl  
36 5 10.0 17 2 US-08-473-656A-20 Sequence 20, Appl  
37 5 10.0 17 2 US-08-473-656A-27 Sequence 27, Appl  
38 5 10.0 17 2 US-08-473-656A-29 Sequence 29, Appl  
39 5 10.0 17 2 US-08-473-656A-30 Sequence 30, Appl  
40 5 10.0 17 2 US-08-473-656A-37 Sequence 37, Appl  
41 5 10.0 17 2 US-08-473-656A-38 Sequence 38, Appl  
42 5 10.0 17 3 US-08-483-931B-12 Sequence 12, Appl  
43 5 10.0 17 3 US-08-483-931B-13 Sequence 13, Appl  
44 5 10.0 17 3 US-08-483-931B-18 Sequence 18, Appl  
45 5 10.0 17 3 US-08-483-931B-19 Sequence 19, Appl

#### ALIGNMENTS

RESULT 1  
US-08-986-837-8  
; Sequence 8, Application US/08986837  
; Patent No. 6221676  
; GENERAL INFORMATION:  
; APPLICANT: Lam, Bing K.  
; APPLICANT: Penrose, John F.  
; APPLICANT: Frank, Austen K.  
; TITLE OF INVENTION: DNA ENCODING HUMAN LEUKOTRIENE C4 SYNTHASE,  
; TITLE OF INVENTION: POLYPEPTIDES, AND USES THEREFOR  
; FILE REFERENCE: Docket No. 6221676 0092662-0019 (BWH269)  
; CURRENT APPLICATION NUMBER: US/08/986,837  
; CURRENT FILING DATE: 1997-12-18  
; EARLIER APPLICATION NUMBER: 08/246,991  
; EARLIER FILING DATE: 1994-05-20  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-986-837-8

Query Match 12.0%; Score 6; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 LRTLLP 24  
Db 8 LRTLLP 13

RESULT 2  
US-08-447-702-1  
; Sequence 1, Application US/08447702  
; Patent No. 5629190  
; GENERAL INFORMATION:  
; APPLICANT: Petre, Dominique  
; APPLICANT: Cerbelesaud, Edith  
; APPLICANT: Levy-Schil, Sophie  
; APPLICANT: Crouzet, Joel  
; TITLE OF INVENTION: Polypeptides Possessing A Nitrilase  
; TITLE OF INVENTION: Activity, DNA Sequence Coding for Said Polypeptides,  
; Patent No. 5629190  
; TITLE OF INVENTION: Expression Cassettes and Host Microorganisms Enabling Them  
; TITLE OF INVENTION: to be Obtained, and Method of Converting Nitriles to Carbox  
; TITLE OF INVENTION: Iates by Means of Said Polypeptides  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:

PT useful in pharmaceuticals, cosmetics, veterinary applications and  
PT agrochemicals  
XX  
XX  
XX Disclosure: Page 38; 55pp; English.  
XX  
CC The present invention relates to a method for identifying  
CC physico-chemical and/or topological parameters associated with biological  
CC activity. The method involves selecting the first subset from  
CC predetermined set of physico-chemical parameters, determining their value  
CC of function, and selecting the second subset from physico-chemical  
CC parameters based on the values of function, such that each second subset  
CC is more closely associated with the activity than the first subset. The  
CC selected physico-chemical parameters are useful for developing criteria  
CC for screening candidate molecules and are suitable for use in silico  
CC screening of compounds. The compounds may be used in pharmaceuticals,  
CC cosmetics, agrochemicals, biomaterials and in veterinary applications. It  
CC is also useful as an antibiotic or antifungal agent. The present  
CC sequence is an immunosuppressive peptide, DK.75-84, identified by in  
CC silico screening. The immunosuppressive activity of the peptide that  
CC prevents allograft rejection is tested in a heterotopic allograft model  
CC of mouse.  
XX  
SQ Sequence 10 AA;

Query Match 10.0%; Score 5; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 19 LRTLL 23  
Db 4 lrtll 8

Search completed: September 15, 2001, 12:56:41  
Job time: 105 sec

XX PS Example 1; Page 19; 41pp; English.

XX CC Peptides AAW33784-98 and AAW33778-9 were assayed for their immunomodulating activity. A peptide-type compound or variant is claimed which has immunomodulating activity, including the N-terminal acylated and/or C-terminal amidated or esterified forms of up to 60 amino acids, and where the peptide-type compound comprises the formula: A-B, where A, B = (R aa76-77L) (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino acid. The sequence in the brackets may optionally be absent or truncated at any peptide type bond within the brackets. The compounds comprise amino acid sequences related to a Class I HLA-B alpha domain (positions 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in a host or in vitro. They can also be used in combination with antigenic peptides or proteins of interest to activate CTLs. They can also inhibit the proliferation of T cells in response to anti-CD3. The peptide can be used for preventing rejection of transplants or for treating autoimmune diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus. The products can also be used for detection and diagnosis.

XX SQ Sequence 10 AA;

Query Match 10.0%; Score 5; DB 19; Length 10;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
| | | | |

DB 4 lrtll 8

RESULT 14  
AAV50273  
ID AAV50273 standard; Peptide; 10 AA.  
XX AC AAV50273;  
DT 12-JAN-2000 (first entry)  
DE Neutrophil-activating pancreatic derived peptide 73.  
XX KW Cell activation; pancreas; treatment; cardiovascular disease; trauma; inflammatory disease; autoimmune diseases; arthritis; diabetes; stroke; organ rejection; ischemia; Alzheimer's disease; myocardial infarction; haemorrhagic shock; diabetic retinopathy; venous insufficiency; angina; trauma; protease inhibitor; hypertension; sepsis.  
XX OS Unidentified.  
XX PN WO9946367-A2.  
XX PD 16-SEP-1999.  
XX PF 11-MAR-1999; 99WO-US05247.  
XX PR 11-MAR-1998; 98US-0038894.  
XX PA (CELL-) CELL ACTIVATION INC.  
XX PA (REGC ) UNIV CALIFORNIA.  
XX PA (SCRI ) SCRIPPS RES INST.  
XX PI Stoughton RB, Schmid-Schonbein GW, Hugli TE, Kistler E;  
XX WPI; 1999-580234/49.  
XX PT Use of cell activating compositions in developing products for diagnosis and treatment of e.g. cardiovascular, inflammatory, autoimmune or Alzheimer's disease, trauma, arthritis, organ rejection, diabetes, stroke or ischemia.

XX PS Example 9; Page 183; 184pp; English.

XX CC This invention describes a novel method for the use and preparation of cell activating compositions which involves preparing a cell activating composition comprising (a) homogenizing pancreatic tissue in buffer at about neutral or higher pH to produce a homogenate; (b) removing particulates from the homogenate; (c) optionally incubating the resulting homogenate, with particulates removed with a protease; and (d) fractionating the homogenate and selecting fractions that exhibit cell activation activity. The methods can be used for improving treatment outcome or reducing risk of treatment of e.g. cardiovascular disease, inflammatory disease, trauma, autoimmune diseases, arthritis, organ rejection, diabetes and diabetic complications, stroke, ischemia, Alzheimer's disease, myocardial infarction, haemorrhagic shock, diabetic retinopathy, diabetes, venous insufficiency, unstable angina or trauma. They can be used in the veterinary treatment of a non-human subject. Protease inhibitors can be used to lower cell activation resulting from these diseases and deficiencies. The detection of an elevated level of hydrogen peroxide can be used to detect an inflammatory condition. An elevated level of hydrogen peroxide in plasma or whole blood and in the presence of superoxide dismutase (SOD) indicates leukocyte up regulation, e.g. indicative of the onset of an acute cardiovascular disorders, such as disease onset or ischemic complications. An elevated level of hydrogen peroxide in plasma or whole blood and a low level in the presence of SOD is indicative of a chronic or immune compromised condition e.g. hypertension or sepsis. AAY50201-Y50334 represent peptides used in the method of the invention.

XX SQ Sequence 10 AA;

Query Match 10.0%; Score 5; DB 20; Length 10;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 SVLEA 39  
| | | | |

DB 1 svlea 5

RESULT 15  
AAV72484  
ID AAY72484 standard; peptide; 10 AA.  
XX AC AAY72484;  
DT 24-APR-2001 (first entry)  
DE Immunosuppressive peptide, Dk.75-84, to prevent allograft rejection.  
XX KW Immunosuppressive; allograft rejection; topological parameter; physico-chemical parameter; in silico screening; pharmaceutical; cosmetic; agrochemical; biomaterial; veterinary application.  
XX OS Unidentified.  
XX PN WO200079263-A2.  
XX PD 28-DEC-2000.  
XX PF 15-MAY-2000; 2000WO-EP04338.  
XX PR 18-JUN-1999; 99EP-0401526.  
XX PA (SYNT-) SYNT:EM SA.  
XX PI Lahana R, Clair P, Yasri A;  
XX WPI; 2001-091623/10.  
XX PT Identifying active candidate molecules on the basis of selected physico-chemical parameters, for in silico screening of compounds

```

PF 05-APR-1996; 96WO-US04710.
XX
PR 12-MAY-1995; 95US-0440504.
XX
XX (SANG-) SANGSTAT MEDICAL CORP.
XX
XX Buelow R;
XX
XX WPI; 1996-518410/51.
XX
XX Treatment of auto-immune disease by admin. of peptide(s) corresp. to
PT major histocompatibility complex antigens - esp. for delaying onset
PT of clinical symptoms of insulin dependent diabetes by modulating T
PT cell mediated attack on target cells
XX
XX Example 1; Page 12; 24pp; English.
XX
XX AA07521-W07524, and AA07527 represent T-cell modulating peptides that
CC can be used in the method of the invention. These sequences are based on
CC a portion of the generic peptide corresponding to residues 70-91 of the
CC alpha1-domain of the major histocompatibility complex (MHC) class I
CC antigen (see AA07510). The method is for affecting the course of an
CC autoimmune disease involving T-cell mediated destruction of tissue in
CC mammals. These peptides are used especially to treat insulin-dependent
CC diabetes mellitus, preferably being administered during the pre-clinical
CC stage to delay onset of the disease. Other diseases that can be treated
CC are multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus
CC vulgaris, Sjogren's disease, thyroid disease, Hashimoto's thyroiditis,
CC myasthenia gravis, etc. The peptides modulate T-cell mediated attack on
CC autologous target cells, and may also reduce inflammation, swelling, and
CC release of cytokines, perforins, granzymes etc. associated with T cell
CC activation.
XX
XX Sequence 10 AA;
SQ
Query Match 10.0%; Score 5; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 19 LRTLL 23
Db 4 lrtll 8
| | | | |
RESULT 12
AA07517
ID AA07517 standard; peptide; 10 AA.
XX
AC AA07517;
XX
XX 04-AUG-1997 (first entry)
XX
XX T-cell modulating peptide #6.
XX
XX T-cell modulator; autoimmune disease; tissue destruction; alpha1-domain;
KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;
KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;
KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;
KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;
KW autologous target cell; cytokine release; T cell activation; therapy.
XX
XX Synthetic.
OS
XX WO9635443-A1.
PN
XX 14-NOV-1996.
XX
XX 05-APR-1996; 96WO-US04710.
XX
XX 12-MAY-1995; 95US-0440504.
XX
XX (SANG-) SANGSTAT MEDICAL CORP.
XX
XX Buelow R;
XX
XX WPI; 1996-518410/51.
XX
XX Treatment of auto-immune disease by admin. of peptide(s) corresp. to
PT major histocompatibility complex antigens - esp. for delaying onset
PT of clinical symptoms of insulin dependent diabetes by modulating T
PT cell mediated attack on target cells
XX
XX Example 1; Page 12; 24pp; English.
XX
XX AA07521-W07524, and AA07527 represent T-cell modulating peptides that
CC can be used in the method of the invention. These sequences are based on
CC a portion of the generic peptide corresponding to residues 70-91 of the
CC alpha1-domain of the major histocompatibility complex (MHC) class I
CC antigen (see AA07510). The method is for affecting the course of an
CC autoimmune disease involving T-cell mediated destruction of tissue in
CC mammals. These peptides are used especially to treat insulin-dependent
CC diabetes mellitus, preferably being administered during the pre-clinical
CC stage to delay onset of the disease. Other diseases that can be treated
CC are multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus
CC vulgaris, Sjogren's disease, thyroid disease, Hashimoto's thyroiditis,
CC myasthenia gravis, etc. The peptides modulate T-cell mediated attack on
CC autologous target cells, and may also reduce inflammation, swelling, and
CC release of cytokines, perforins, granzymes etc. associated with T cell
CC activation.
XX
XX Sequence 10 AA;
SQ
Query Match 10.0%; Score 5; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 19 LRTLL 23
Db 4 lrtll 8
| | | | |
RESULT 13
AAW33785
ID AAW33785 standard; peptide; 10 AA.
XX
AC AAW33785;
XX
XX 19-JUN-1998 (first entry)
XX
XX Peptide B2705.75-84 tested for immunomodulating activity.
XX
XX Immunomodulating dimer; immunosuppressant drug; CTL activation;
KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
KW rejection.
XX
XX Synthetic.
OS
XX Homo sapiens.
XX
XX WO9744351-A1.
PN
XX 27-NOV-1997.
XX
XX 22-MAY-1997; 97WO-US08689.
XX
XX 24-MAY-1996; 96US-0653294.
XX
XX (STPD ) UNIV LELAND STANFORD JUNIOR.
XX
XX Buelow R, Clayberger C, Krensky AM;
XX
XX WPI; 1998-086530/08.
XX
XX New immunomodulating dimer peptide(s) - based on a Class I HLA-B
PT alpha-1 domain, used for preventing rejection of transplants or
PT treating autoimmune diseases

```



XX PS Example 26; Page 27; 47pp; English.

CC The patent discloses dynorphin-like polypeptides which are shorter

CC than dynorphin (7 to 9 amino acids) and which contain either a D-

CC amino acid and/or an N-alkyl deriv. of an L- or D-amino acid to

CC provide in-vivo stability upon intravenous administration. The

CC peptides have a much greater analgesic effect than dynorphin

CC due to their greater stability in the blood (e.g. IC50 values can be

CC as low as 0.04 nM when tested by the rabbit vas deferens method; cf.

CC 17.4 nM for dynorphin).

CC The present peptide is a specific example of the peptides.

XX SQ Sequence 9 AA;

Query Match 10.0%; Score 5; DB 15; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRER 11

Db 5 Lrrer 9

|||||

RESULT 7

AAR84869

ID AAR84869 standard; peptide; 9 AA.

AC AAR84869;

XX 30-MAY-1996 (first entry)

DT 30-MAY-1996 (first entry)

XX H2Kb 75-83 immunogenic peptide.

DE Antigen: epitope; cell mediated; immune specific; cancer;

XX infection; infestation; mucin-1; MUC-1; tumour; H2Kb 75-83;

KW immunogenic peptide.

XX Synthetic.

XX WO9527505-A1.

XX 19-OCT-1995.

PD 12-APR-1994; 94US-0229606.

XX 12-APR-1994; 95WO-US04540.

XX (BIOM-) BIOMIRA INC.

XX Ding L, Koganty RR, Longenecker BM, Reddish MA;

PI WPI; 1995-373528/48.

DR New cell-mediated immune-specific immunogenic compsns. - used in

XX prophylaxis and treatment of cancer, microbial infections, viral

PT infections and parasitic infestations

XX Example 1; Page 62; 141pp; English.

XX A claimed cell-mediated immune (CMI)-specific immunogenic compsn.

CC comprises a conjugate of a primary antigen bearing a primary epitope,

CC with an immunomodulatory peptide (IP), i.e. the present peptide.

CC The IP comprises an allopeptide moiety of at least 5 amino acids,

CC whose sequence corresponds to a polymorphic region of a MHC

CC encoded polymorphic Class I or II antigen. The compsn. can be

CC used to elicit a CMI-specific response which is prophylactic, or

CC therapeutic for, e.g. microbial and viral infections, parasitic

CC infestations and cancer, partic. MUC-1 expressing tumour cells

CC when the present peptide is the IP, and a MUC-1 epitope is the

CC primary epitope.

SQ Sequence 9 AA;

Query Match 10.0%; Score 5; DB 16; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23

Db 4 lrtll 8

|||||

RESULT 8

AAR41212

ID AAR41212 standard; peptide; 10 AA.

XX AAR41212;

AC 15-MAR-1994 (first entry)

DT 15-MAR-1994 (first entry)

XX Peptide fragment of Class I HLA peptide.

DE Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;

XX parasitic disease; cytotoxic T lymphocyte; modulation.

KW Synthetic.

XX OS WO9317699-A.

XX PN 16-SEP-1993.

PD 25-FEB-1993; 93WO-US01758.

XX 02-MAR-1992; 92US-0844716.

PR (STRD ) UNIV LELAND STANFORD JUNIOR.

XX Clayberger CA, Krensky AM;

PI WPI; 1993-303134/38.

DR New peptide(s) based on Class I HLA antigen domains - used for

XX modulating cytotoxic T-lymphocyte activity towards targets

PT Claim 11; Page 54; 61pp; English.

XX The peptide is used to modulate cytotoxic T-lymphocyte (CTL)

CC activity, either by inhibition or stimulation. It can be used

CC for inhibiting CTL toxicity in transplantations, for inducing CTL

CC activity in parasitic diseases and neoplasia and in studies on viral

CC infection. The peptide can also be used for identifying CTLs which

CC bind to it and removing subsets of CTLs from a T-cell composition.

CC This peptide sequence is more commonly found within larger peptide

CC compounds of not more than 30 amino acids in length.

XX SQ Sequence 10 AA;

Query Match 10.0%; Score 5; DB 14; Length 10;

Best Local Similarity 100.0%; Pred. No. 93;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23

Db 4 lrtll 8

|||||

RESULT 9

AAR95423

ID AAR95423 standard; peptide; 10 AA.

XX AC AAR95423;

XX

RESULT 4  
 AAR50267  
 ID AAR50267 standard; peptide; 6 AA.  
 XX  
 AC AAR50267;  
 DT 13-OCT-1994 (first entry)  
 XX  
 DE Peptide corresponding to a shared sequence of an HLA molecule.  
 XX  
 KW Peptide; HLA; human leucocyte antigen; diagnosis; treatment;  
 KW autoimmune disease; uveitis; spondylitis; psoriasis;  
 KW inflammatory bowel disease; enteric bacteria;  
 KW Salmonella typhimurium.  
 XX  
 OS Salmonella typhimurium.  
 XX  
 PN W09405303-A.  
 XX  
 PD 17-MAR-1994.  
 XX  
 PF 31-AUG-1993; 93WO-US08214.  
 XX  
 PR 31-AUG-1992; 92US-0944143.  
 XX  
 PA (OKLA ) UNIV OKLAHOMA STATE.  
 XX  
 PI Harley JB, Scofield RH;  
 XX  
 DR WPI; 1994-100843/12.  
 XX  
 PT Peptide corresponding to shared sequences of HLA molecules and  
 PT enteric bacteria - used for the diagnosis and treatment of  
 PT auto-immune disorders, partic. spondylarthropathies.  
 XX  
 PS Disclosure; Page 15; 58pp; English.  
 XX  
 CC The peptide is a fragment of the HLA B27 hypervariable region  
 CC described in AAR50266. The peptide can be used for the diagnosis and  
 CC treatment of autoimmune disorders, e.g. spondyloarthropies including  
 CC uveitis and spondylitis associated with inflammatory bowel disease  
 CC or psoriasis.  
 XX  
 SQ Sequence 6 AA;  
  
 Query Match 10.0%; Score 5; DB 15; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 19 LRTLL 23  
 Db 2 lrtll 6  
  
 RESULT 5  
 AAR59948  
 ID AAR59948 standard; peptide; 6 AA.  
 XX  
 AC AAR59948;  
 XX  
 DT 14-FEB-1995 (first entry)  
 XX  
 DE Peptide signal sequence for treating E. coli infections.  
 XX  
 KW Therapeutic; metabolic interactions; PSS; analogues.  
 XX  
 OS Synthetic.  
 XX  
 PN W09416328-A.  
 XX  
 PD 21-JUL-1994.

XX 30-DEC-1993; 93WO-US12679.  
 PF  
 XX 30-DEC-1992; 92US-0997727.  
 PR  
 XX (RATH/) RATH M.  
 PA  
 XX Rath M.  
 PI  
 XX WPI; 1994-249399/30.  
 DR  
 XX Identifying peptide signal sequences in a protein - and use of  
 PT their synthetic analogues for treating or preventing, e.g.  
 PT cardiovascular and auto-immune disease, infections and cancer.  
 XX  
 PS Claim 18; Page 10; 28pp; English.  
 XX  
 CC The sequence is that of a peptide signal sequence which can be used  
 CC to treat E. coli infections.  
 CC See also AAR59944-83.  
 XX  
 SQ Sequence 6 AA;  
  
 Query Match 10.0%; Score 5; DB 15; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 2 SSKEK 6  
 Db 2 sskek 6  
  
 RESULT 6  
 AAR61019  
 ID AAR61019 standard; peptide; 9 AA.  
 XX  
 AC AAR61019;  
 XX  
 DT 23-APR-1995 (first entry)  
 XX  
 DE Dynorphin-like polypeptide.  
 XX  
 KW Dynorphin; opioid analgesic; stable; stability; intravenous.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "Metyr"  
 FT Misc-difference 8 /note= "D-Glu"  
 FT Modified-site 9 /note= "Arg-NH2"  
 FT  
 XX EP614913-A.  
 PN  
 XX 14-SEP-1994.  
 PD  
 XX 08-NOV-1985; 94EP-0107769.  
 XX  
 PF 09-NOV-1984; 84JP-0236076.  
 XX  
 PR (EISA ) EISAI CO LTD.  
 XX  
 PA Arakawa Y, Araki S, Ikeda M, Kaneko T, Nakazawa T;  
 XX Tachibana S, Tsuchiya Y, Yamatsu K, Yoshino H;  
 PI  
 XX WPI; 1994-281157/35.  
 DR  
 XX New dynorphin polypeptide derivs. for use as analgesics - contg.  
 PT a D-amino acid or an N-alkyl deriv. of a D- or L-amino acid for  
 PT in vivo stability.



PT Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -

XX Example 5; Page 36; 62pp; English.

XX The present sequence is that of a peptide derived from human  
 CC PHELIIX (see AAY79269), a novel transcription factor that is  
 CC normally expressed only in testis tissue, but which is up-regulated  
 CC in prostate and some other cancers. The peptide was conjugated to  
 CC keyhole limpet haemocyanin and used to raise polyclonal antiserum  
 CC in rabbit. The antiserum demonstrated specificity for PHELIIX and  
 CC may therefore be useful for assessing the expression of PHELIIX in  
 CC patient samples.

XX Sequence 15 AA;

Query Match 28.0%; Score 14; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-08;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HSSKEKLRRIKY 14  
 |||||  
 DB 1 hsskeklrreriky 14

RESULT 2

AAB15385  
 ID AAB15385 standard; Protein; 27 AA.

XX AC AAB15385;

XX DT 06-DEC-2000 (first entry)

XX DE Comamonas testosteroni nitrilase protein N-terminus.

XX KW Nitrilase; hydrolysis; dinitrile; carboxylic acid; adiponitrile;  
 KW adipic acid; nylon 6,6; 5-cyanovaleic acid.

XX OS Comamonas testosteroni.

XX PN FR2694571-A1.

XX PD 11-FEB-1994.

XX PF 10-AUG-1992; 92FR-0009882.

XX PR 10-AUG-1992; 92FR-0009882.

XX PA (RHON) RHONE POULENC CHIM.

XX PI Petre D, Cerebelaud E, Levy-Schil S, Crouzet J;

XX DR WPT; 1994-076687/10.

XX PT New gene for nitrilase from Comamonas testosteroni - esp. for  
 PT converting di-nitrile(s) to di-carboxylic acids, also the new enzyme  
 PT and transformed microorganisms

XX PS Example 1; Page 10; 33pp; French.

XX This sequence represents the N-terminal sequence of a nitrilase, able  
 CC to hydrolyse nitriles to carboxylic acids, isolated from the bacterium  
 CC Comamonas testosteroni. The N-terminal sequence was obtained by peptide  
 CC sequencing of the purified enzyme. The nitrilase, or microorganisms  
 CC which produce it, is useful for converting dinitriles NC.R-CN (R=1-10C  
 CC alkylene) to corresponding acids, e.g. adiponitrile to adipic acid (for  
 CC nylon 6,6 manufacture) or 5-cyanovaleic acid (as their NH4 salts). The  
 CC protein provides rapid and complete hydrolysis of dinitriles; contrast  
 CC known nitrilases with which hydrolysis of the second CN is usually very  
 CC slow.

SQ Sequence 27 AA;

Query Match 12.0%; Score 6; DB 15; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 37 LEATVD 42  
 |||||  
 DB 20 leatvd 25

RESULT 3

AAP30295  
 ID AAP30295 standard; Protein; 6 AA.

XX AC AAP30295;

XX DT 20-APR-1992 (first entry)

XX DE Sequence of AAs 26-31 of the E. coli heat stable toxin  
 DE which determine an H-epitope.

XX KW Synthetic vaccine; antigen; allergen; immunological response;  
 KW antibody.

XX OS Escherichia coli.

XX PN EP93851-A.

XX PD 16-NOV-1983.

XX PF 11-MAR-1983; 83EP-0102392.

XX PR 15-MAR-1982; 82US-0358150.

XX PR 28-JAN-1983; 83US-0461802.

XX PR 12-JUN-1981; 81US-0272855.

XX PR 09-JAN-1981; 81US-0223558.

XX PR 16-DEC-1986; 86US-0942562.

XX PA (NYBL-) NEW YORK BLOOD CENT.

XX PI Hopp TP;  
 DR WPT; 1983-822049/47.

XX PT Synthetic vaccine - contains peptide residue coupled to higher  
 PT alkyl or alkenyl Gps. and with 6 amino acids in residue

XX PS Claim 50; Page 46; 54pp; English.

XX The inventors claim a synthetic vaccine which comprises a peptide  
 CC residue coupled to an alkyl or alkenyl gp. having at least 12C or  
 CC other lipophilic substance. The residue contains a sequence of 6 AAs  
 CC corresp. to the SQ of such AAs in a protein antigen or allergen where  
 CC the greatest local average hydrophobicity is found. Pref. the AAs in  
 CC the peptide do not exceed 50 residues, and they esp. contain 12-18  
 CC residues. The alkyl or alkenyl gp. pref. contains 12-24C and it is  
 CC pref. coupled to the terminal amino gp. of the residue opt. via a CO  
 CC gp. if a lipophilic substance is used, it is pref. palmitic, stearic,  
 CC behenic, oleic or mycolic acid.

XX SQ Sequence 6 AA;

Query Match 10.0%; Score 5; DB 4; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SSKEK 6

DB 2 sskek 6

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:54:56 ; Search time 17.5 seconds  
(without alignments)  
173.211 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_189  
Perfect score: 50  
Sequence: 1 HSSKEKLRRRIKYCQQLR.....NDAASVLEATVDYVKYIREK 50

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 217092

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_0601.\*

- 1: /cgnl\_9/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /cgnl\_9/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /cgnl\_9/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /cgnl\_9/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /cgnl\_9/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /cgnl\_9/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /cgnl\_9/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /cgnl\_9/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /cgnl\_9/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /cgnl\_9/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /cgnl\_9/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /cgnl\_9/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /cgnl\_9/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /cgnl\_9/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /cgnl\_9/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /cgnl\_9/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /cgnl\_9/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /cgnl\_9/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /cgnl\_9/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /cgnl\_9/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /cgnl\_9/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	14	28.0	15	AA199270	PHLIX peptide use
2	6	12.0	27	AA15385	Comamonas testoste
3	5	10.0	6	AA130295	Sequence of AAS 26
4	5	10.0	6	AA150267	Peptide correspond
5	5	10.0	6	AA159948	Peptide signal seq
6	5	10.0	9	AA161019	Dynorphin-like pol
7	5	10.0	9	AA184869	H2kb 75-83 immunog
8	5	10.0	10	AA141212	Peptide fragment o
9	5	10.0	10	AA195423	HLA-B2705.75-84.
10	5	10.0	10	AA183075	HLA-B2702 CTL modu
11	5	10.0	10	AA17524	T-cell modulating

T-cell modulating  
Peptide B2705.75-8  
Neutrophil-activat  
Immunosuppressive  
Glucose transport  
Guanine thymine bi  
Human Class I HLA-  
Human Class I HLA-  
HIV principal neut  
CAEV env gene TM1  
Oestrogen receptor  
Peptide fragment d  
HLA B27 hypervaria  
Glucose transport  
Human MHC 1 and HL  
Human HLA-B27-(G2-  
Human [Phe74]-HLA-  
Human MHC 1 alpha  
Human MHC 1 alpha  
Human MHC 1 alpha  
Human MHC 1 alpha  
Human MHC 1 alpha  
Glucose transport  
MHC peptide repeat  
MHC peptide repeat  
Peptide membrane b  
P/- swap 2 mutain o  
Bacillus amyloliqu  
UL snRNP 70K prote  
Human neuroendocri  
Human MHC alpha-1  
Human MHC Class I-  
Human MHC 1 alpha

#### ALIGNMENTS

RESULT 1  
AA199270  
ID AA199270 standard; Peptide; 15 AA.  
XX  
AC AA199270;  
XX  
DT 03-JUL-2000 (first entry)  
XX  
DE PHLIX peptide used to raise antibody.  
XX  
KW PHLIX; human; testis-specific; transcription factor;  
KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
KW therapy; diagnosis; vaccine; antibody.  
XX  
OS Homo sapiens.  
XX  
PN WO200012709-A2.  
XX  
PD 09-MAR-2000.  
XX  
PF 31-AUG-1999; 99WO-US20137.  
XX  
PR 31-AUG-1998; 98US-0098610.  
PR 31-OCT-1998; 98US-0106524.  
XX  
(UROC-) UROGENESYS INC.  
PA (AFAR/) AFAR D E.  
PA (HUBE/) HUBERT R S.  
PA (RAIT/) RAITANO A B.  
XX  
Afar DE, Hubert RS, Raitano AB;  
PI WPI; 2000-237872/20.  
XX



Query Match 12.5%; Score 7; DB 14; Length 331;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YCCEQLR 26  
 | | | | |  
 DB 127 YCCEQLR 133

RESULT 13  
 Q9HV76 PRELIMINARY; PRT; 390 AA.  
 AC Q9HV76;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE PROBABLE AMINOTRANSFERASE.  
 GN PA4722.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Phan X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,  
 RA Hickey M.J., Brinkman F.S.D., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.,  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 DR EMBL; AE004885; AAG08108.1; -  
 DR InterPro; IPR001176; -  
 DR InterPro; IPR001511; -  
 DR Pfam; PF00155; aminotran\_1;  
 DR PRINTS; PR00753; ACCSYNTHASE.  
 DR PROSITE; PS00105; AA\_TRANSFER\_CLASS\_1; UNKNOWN\_1.  
 KW Transferase; Aminotransferase.  
 SQ SEQUENCE 390 AA; 42567 MW; CB6A8E94B5C674C3 CRC64;

Query Match 12.5%; Score 7; DB 2; Length 390;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 DAASVLE 44  
 | | | | |  
 DB 215 DAASVLE 221

RESULT 14  
 Q9H553 PRELIMINARY; PRT; 416 AA.  
 AC Q9H553;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE BA13B9.1 (NOVEL PROTEIN SIMILAR TO A GLYCOSYLTRANSFERASE).  
 GN BA13B9.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Babbage A.;  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL137067; CAC07999.1; -.

KW Transferase.  
 SQ SEQUENCE 416 AA; 47091 MW; 778DB1FD069E7F29 CRC64;

Query Match 12.5%; Score 7; DB 4; Length 416;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KISLLHS 8  
 | | | | |  
 DB 308 KISLLHS 314

RESULT 15  
 Q9LYR0 PRELIMINARY; PRT; 443 AA.  
 ID Q9LYR0;  
 AC Q9LYR0;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE HYPOTHETICAL 50.8 KDA PROTEIN.  
 GN T2N19\_110.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; euroids II;  
 OC Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Dueterhoeft A.,  
 RA Bancroft I., Meves H.W., Rudd S., Lemcke K., Mayer K.F.X.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL163572; CAB87153.1; -  
 DR InterPro; IPR000048; -  
 DR Pfam; PF00612; IQ; 2;  
 DR SMART; SM00015; IQ; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 443 AA; 50820 MW; EA3C8B85E2041D83 CRC64;

Query Match 12.5%; Score 7; DB 10; Length 443;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRRERIK 19  
 | | | | |  
 DB 227 LRRERIK 233

Search completed: September 15, 2001, 12:50:05  
 Job time: 265 sec

Q9L8Y3 PRELIMINARY; PRT; 268 AA.  
AC Q9L8Y3;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE VANYB (D-D-CARBOXYPEPTIDASE).  
GN VANYB OR VANYB2.  
OS Enterococcus faecalis (Streptococcus faecalis), and  
OS Enterococcus faecium (Streptococcus faecium).  
OG Plasmid pIP834.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;  
OC Enterococcus.  
OC NCBI\_TaxID=1351, 1352;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.faecalis; STRAIN=BM4382; PLASMID=PIP834; TRANSPONSON=TN1549;  
RA Garnier F., Taourit S., Glaser P., Courvalin P., Gallmand M.;  
RT "Characterization of transposon Tn1549 conferring VanB-type resistance  
in Enterococcus sp."  
RL Microbiology 0:0-0(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=E.faecium; STRAIN=TSGH1;  
RA Lu J.-J., Perng C.-L., Ho M.-F., Lu C.-L.;  
RT "Whole sequence of vanB gene clusters of Enterococcus faecium  
TSGH1."  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF192329; AAF72360.1; -;  
DR EMBL; AF310956; AAG34688.1; -;  
KW Plasmid; Carboxypeptidase.  
SQ SEQUENCE 268 AA; 30424 MW; 53FFBC8BDE3B30B1 CRC64;

Query Match 12.5%; Score 7; DB 2; Length 268;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 ASVLEAT 46  
| | | | |  
Db 46 ASVLEAT 52

RESULT 10  
083552  
ID 083552 PRELIMINARY; PRT; 319 AA.  
AC 083552;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE GTP-BINDING PROTEIN (ERA).  
GN TP0541.  
OS Treponema pallidum.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
OC NCBI\_TaxID=160;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NICHOLS;  
RX MEDLINE=98332770; PubMed=9665876;  
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
RA McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
RA Venter J.C.;  
RT "Complete genome sequence of Treponema pallidum, the syphilis  
spirochete."  
RL Science 281:375-388(1998).  
DR EMBL; AE001229; AAC65525.1; -;  
DR HSSP; P06616; 1EGA.  
DR TIGR; TP0541; -;  
SQ SEQUENCE 319 AA; 35742 MW; 3DCED9411D1F261A CRC64;

Query Match 12.5%; Score 7; DB 2; Length 319;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLLPY 31  
| | | | |  
Db 287 LRTLLPY 293

RESULT 11  
Q9Y960  
ID Q9Y960 PRELIMINARY; PRT; 326 AA.  
AC Q9Y960;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL 35.1 KDA PROTEIN APE2425.  
GN APE2425.  
OS Aeropyrum pernix.  
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococaceae;  
OC Aeropyrum.  
OC NCBI\_TaxID=56636;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K1;  
RX MEDLINE=99310339; PubMed=10382966;  
RA Kavarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,  
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,  
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
RT "Complete genome sequence of an aerobic hyper-thermophilic  
crenarchaeon, Aeropyrum pernix K1."  
RL DNA Res. 6:83-101(1999).  
DR EMBL; AP000064; BAA81440.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 326 AA; 35093 MW; 909665F6372DD04C CRC64;

Query Match 12.5%; Score 7; DB 1; Length 326;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVLEA 45  
| | | | |  
Db 292 AASVLEA 298

RESULT 12  
O11316  
ID O11316 PRELIMINARY; PRT; 331 AA.  
AC O11316;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)  
DE CLONE B1-13 HOMOLOG OF VACCINIA E2L (B1-13) (FRAGMENT).  
GN B1-13.  
OS Molluscum contagiosum virus subtype 1 (NCV1).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Molluscipoxvirus.  
OC NCBI\_TaxID=10280;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,  
RA Lopez-Esteban J.L., Esteban M., Martin-Gallardo A.;  
RL Virus Genes 0:0-0(0).  
DR EMBL; U86899; AAB57946.1; -;  
FT NON\_TER 1  
FT NON\_TER 331 331  
SQ SEQUENCE 331 AA; 37009 MW; E776B399590B997E CRC64;

```

RP SEQUENCE FROM N.A.
RA Dimons G., Groenendijk J., Wijbrandi J., Reijans M., Groenen J.,
RA Diergaarde P., Van der Lee T., Bleeker M., Onstenk J., De Both M.,
RA Haring M., Mes J., Cornelissen B., Zabeau M., Vos P.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF118127; AAD27815.1; -.
DR InterPro; IPR000767; -.
DR InterPro; IPR001593; -.
DR InterPro; IPR001611; -.
DR InterPro; IPR002182; -.
DR Pfam; PF00560; LRR; 3.
DR Pfam; PF00931; NB-ARC; 1.
DR PRINTS; PR00364; DISEASERST.
DR ProDom; PD003035; -. 1.
SQ SEQUENCE 1266 AA; 144826 MW; 8392FFBE7FED7FD5 CRC64;

Query Match 14.3%; Score 8; DB 10; Length 1266;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EQLRTLLP 30
DB 547 EQLRTLLP 554
|||||

RESULT 6
Q9EPS1 PRELIMINARY; PRT; 108 AA.
AC Q9EPS1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE PHOSPHODIESTERASE 1A (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WISTAR; TISSUE-BRAIN;
RA Prime G.R., Sutor B.;
RT "Phosphodiesterase 1A (PDE1A) in rat brain.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327836; AAG48734.1; -.
FT NON_TER 108
SQ SEQUENCE 108 AA; 12437 MW; 5FD104D5B0CF9D8F CRC64;

Query Match 12.5%; Score 7; DB 11; Length 108;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AASVLEA 45
DB 59 AASVLEA 65
|||||

RESULT 7
O07252 PRELIMINARY; PRT; 200 AA.
AC O07252;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE HYPOTHETICAL 21.4 KDA PROTEIN.
GN RV0328 OR MTCV63.33.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekai F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544 (1998).
CC -!- SIMILARITY: TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.
DR EMBL; Z96800; CAB09609.1; -.
DR Tuberculist; RV0328; -.
DR InterPro; IPR001647; -.
DR Pfam; PF00440; tetr; 1.
DR PRINTS; PR00455; HTHTEWR.
KW DNA-binding; Hypothetical protein; Transcription regulation.
SQ SEQUENCE 200 AA; 21411 MW; 81ADC84B554B5EC6 CRC64;

Query Match 12.5%; Score 7; DB 2; Length 200;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VLEATVD 48
DB 88 VLEATVD 94
|||||

RESULT 8
Q9MOV5 PRELIMINARY; PRT; 250 AA.
AC Q9MOV5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE HYPOTHETICAL 27.6 KDA PROTEIN.
GN AT4G05400.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161503; CAB81082.1; -.
KW Hypothetical protein.
SQ SEQUENCE 250 AA; 27627 MW; 7AAA04FBF0A6F707 CRC64;

Query Match 12.5%; Score 7; DB 10; Length 250;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKELR 14
DB 244 SSKELR 250
|||||

RESULT 9
Q9L8Y3

```

```

GN DRC0037.
OS Deinococcus radiodurans.
OC Plasmid Cpl.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.P., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.-J., Lam P., McDonald L., Otterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RT Science 286:1571-1577(1999).
DR EMBL; AE001827; AAF12674.1; -.
DR TIGR; DRC0037; -.
KW Plasmid.
SQ SEQUENCE 703 AA; 77447 MW; 6ED24979D0ECD646 CRC64;

Query Match 14.3%; Score 8; DB 2; Length 703;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AASVLEAT 46
Db 662 AASVLEAT 669
|||||

RESULT 3
O24015
ID O24015 PRELIMINARY; PRT; 1220 AA.
AC O24015;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE RESISTANCE COMPLEX PROTEIN I2C-1.
GN I2C-1.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97290204; PubMed=9144960;
RA Ori N., Eshed Y., Paran I., Presting G., Aviv D., Tanksley S.,
RA Zamir D., Fluhr R.;
RT "The I2C family from the wilt disease resistance locus I2 belongs to
RT the nucleotide binding, leucine-rich repeat superfamily of plant
RT resistance genes."
RL Plant Cell 9:521-532(1997).
DR EMBL; AF004879; AAB63275.1; -.
DR Mendel; 24473; Lyces:3172;24473.
DR InterPro; IPR001611; -.
DR InterPro; IPR002182; -.
DR InterPro; IPR003592; -.
DR Pfam; PF00560; LRR; 4.
DR Pfam; PF00931; NB-ARC; 1.
DR SMART; SM00370; LRR; 1.
DR SMART; SM00370; LRR; 1.
SQ SEQUENCE 1220 AA; 139386 MW; B5997280C8C5EB20 CRC64;

Query Match 14.3%; Score 8; DB 10; Length 1220;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EQLRTLTP 30
Db 545 EQLRTLTP 552
|||||

RESULT 5
O9XET3
ID O9XET3 PRELIMINARY; PRT; 1266 AA.
AC O9XET3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE DISEASE RESISTANCE PROTEIN I2.
GN I2.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98299845; PubMed=9634592;
RA Simons G., Groenendijk J., Wijbrandi J., Reijans M., Groenen J.,
RA Diergaarde P., Van der Lee T., Bleeker M., Onstenk J., De Both M.,
RA Haring M., Mes J., Cornelissen B., Zabeau M., Vos P.;
RT "Dissection of the fusarium I2 gene cluster in tomato reveals six
RT homologs and one active gene copy."
RL Plant Cell 10:1055-1068(1998).
RN [2]

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:50:02 ; Search time 69.13 Seconds  
(without alignments)  
107.176 Million cell updates/sec

Title: us-09-389-000-2\_copy\_134\_189

Perfect score: 56

Sequence: 1 KKLISLLHSSKEKLRRRIKY.....NDAASVLEATVDYVKYIREK 56

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL16:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_unclassified:\*
- 13: sp\_vertebrate:\*
- 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	78.6	425	Q9NX45	Q9NX45 homo sapien
2	8	14.3	703	Q9RZF3	Q9RZF3 deinococcus
3	8	14.3	1220	Q24015	Q24015 lycopersico
4	5	14.3	1240	Q24016	Q24016 lycopersico
5	8	14.3	1266	Q9XET3	Q9XET3 lycopersico
6	7	12.5	108	Q9EPS1	Q9EPS1 rattus norv
7	7	12.5	200	O07252	O07252 mycobacteri
8	7	12.5	250	Q9MOV5	Q9MOV5 enterococcu
9	7	12.5	268	Q9L8Y3	Q9L8Y3 enterococcu
10	7	12.5	319	O83552	O83552 treponema p
11	7	12.5	326	Q9Y960	Q9Y960 aeropyrum p
12	7	12.5	331	O11316	O11316 molluscum c
13	7	12.5	390	Q9HV76	Q9HV76 pseudomonas
14	7	12.5	416	Q9H553	Q9H553 homo sapien
15	7	12.5	443	Q9LYR0	Q9LYR0 arabidopsis
16	7	12.5	500	O86057	O86057 herbaspiril
17	7	12.5	502	Q9ZB31	Q9ZB31 brucella ab
18	7	12.5	514	O28063	O28063 bos taurus
19	7	12.5	519	Q9UFX3	Q9UFX3 homo sapien

20	7	12.5	542	11	Q9EPR9	Q9EPR9 rattus norv
21	7	12.5	549	8	Q9T6Y9	Q9T6Y9 spathiphyll
22	7	12.5	559	8	Q9T6B9	Q9T6B9 ascarina sp
23	7	12.5	561	8	Q9T6T8	Q9T6T8 doryphora s
24	7	12.5	562	8	Q9T6T6	Q9T6T6 peumus bold
25	7	12.5	562	8	Q9T6T0	Q9T6T0 akabia quin
26	7	12.5	565	8	Q9T6U6	Q9T6U6 trochodendr
27	7	12.5	567	8	Q9T701	Q9T701 thottea tom
28	7	12.5	577	8	Q9T6W3	Q9T6W3 liiodendro
29	7	12.5	579	8	Q9T6V8	Q9T6V8 idiospermum
30	7	12.5	580	8	Q9T6V1	Q9T6V1 euptelea po
31	7	12.5	581	8	Q9T6X5	Q9T6X5 podophyllum
32	7	12.5	584	8	Q9T6X3	Q9T6X3 polyalthia
33	7	12.5	586	2	Q9ZDF1	Q9ZDF1 rickettsia
34	7	12.5	586	8	Q9T6X8	Q9T6X8 cinnaomend
35	7	12.5	586	8	Q9T6X0	Q9T6X0 knema later
36	7	12.5	586	8	Q9T6W1	Q9T6W1 cocculus tr
37	7	12.5	588	8	Q9T6Y0	Q9T6Y0 hedyosmum a
38	7	12.5	589	8	Q9T6X2	Q9T6X2 asimina tri
39	7	12.5	589	8	Q9T6A6	Q9T6A6 dicentra sp
40	7	12.5	590	8	Q9T6V9	Q9T6V9 chimonanthu
41	7	12.5	590	8	Q9T6U9	Q9T6U9 lardizabala
42	7	12.5	590	8	Q9T6U3	Q9T6U3 nelumbo nuc
43	7	12.5	591	8	Q9T6Y7	Q9T6Y7 saururus ce
44	7	12.5	591	8	Q9T6X7	Q9T6X7 xanthorhiza
45	7	12.5	591	8	Q9T6U2	Q9T6U2 cinnanomum

#### ALIGNMENTS

```

RESULT 1
Q9NX45 PRELIMINARY: PRT: 425 AA.
ID Q9NX45
AC Q9NX45;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CDNA FLJ20449 FIS, CLONE KAT05575.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000456; BAA91175.1; -
DR InterPro; IPR001092; -
DR SMART; SM00353; HLH; 1.
SQ SEQUENCE 425 AA; 46922 MW; 777E57231F9D6BDD CRC64;

```

Query Match 78.6%; Score 44; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 1.1e-37;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 13 LRRERIKYCCQLRLLPYVKGKNDAAASVLEATVDYVKYIREK 56
      |||
Db 212 LRRERIKYCCQLRLLPYVKGKNDAAASVLEATVDYVKYIREK 255

```

```

RESULT 2
Q9RZF3 PRELIMINARY: PRT: 703 AA.
ID Q9RZF3
AC Q9RZF3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE MODULATION PROTEIN-RELATED PROTEIN.

```



Search completed: September 15, 2001, 12:50:40  
Job time: 285 sec

---

```
Query Match      10.7%; Score 6; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRRERI 18
Db 14 LRRERI 19

RESULT 14
YB4C_SCHPO STANDARD; PRT; 93 AA.
ID O14358;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL 10.7 KDA PROTEIN C30D10.12C IN CHROMOSOME II.
GN SPBC30D10.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajandream M.A., Barrell B.G., Lauber J., Hilbert H.,
RA Duesterhoeft A.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST YGR215W.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 297992; CAB10807.1; -
KW Hypothetical protein.
SQ SEQUENCE 93 AA; 10698 MW; 6ECC74F169DA747B CRC64;

Query Match      10.7%; Score 6; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 QLRTLL 29
Db 55 QLRTLL 60

RESULT 15
ASFP_BOVIN STANDARD; PRT; 134 AA.
ID ASFP_BOVIN
AC P29392;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ACIDIC SEMINAL FLUID PROTEIN PRECURSOR (ASFP).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seminal vesicle;
RX MEDLINE=92181448; PubMed=1543494;
RA Wempe F., Einspanier R., Scheit K.H.;
RT "Characterization by cDNA cloning of the mRNA of a new growth factor
RT from bovine seminal plasma: acidic seminal fluid protein.";
```

```
RL Biochem. Biophys. Res. Commun. 183:232-237(1992).
RN [2]
RP SEQUENCE OF 21-43.
RC TISSUE=Seminal vesicle;
RX MEDLINE=91378963; PubMed=1898381;
RA Einspanier R., Einspanier A., Wempe F., Scheit K.H.;
RT "Characterization of a new bioactive protein from bovine seminal
RT fluid.";
RL Biochem. Biophys. Res. Commun. 179:1006-1010(1991).
RN [3]
RP DISULFIDE BONDS.
RX MEDLINE=94237283; PubMed=8181566;
RA Einspanier R., Krause I., Calvete J.J., Toepfer-Petersen E.,
RA Klostermeyer H., Karg H.;
RT "Bovine seminal plasma ASFP: localization of disulfide bridges and
RT detection of three different isoelectric forms.";
RL FEBS Lett. 344:61-64(1994).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=97475216; PubMed=9334740;
RA Romero A., Romao M.J., Varela P.F., Koelln I., Dias J.M.,
RA Carvalho A.L., Sanz L., Toepfer-Petersen E., Calvete J.J.;
RT "The crystal structures of two spermadhesins reveal the CUB domain
RT fold.";
RL Nat. Struct. Biol. 4:783-788(1997).
CC -!- FUNCTION: STIMULATES CELL DIVISION AND PROGESTERONE SECRETION
CC OF BOVINE GRANULOSA CELLS IN VITRO IN A POTENT AND DOSE DEPENDENT
CC MANNER. THIS PROTEIN APPEARS TO BE A POTENT GROWTH FACTOR WITH
CC EFFECTS ON OVARIAN GRANULOSA CELLS.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- TISSUE SPECIFICITY: SEMINAL VESICLE TISSUE, AMPULLA AND WEAKLY
CC -!- IN TISSUE OF EPIDIDYMIS.
CC -!- SIMILARITY: CONTAINS 1 CUB DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE SPERMADHESIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84603; AAA30745.1; -
DR PIR; PH0213; PH0213;
DR PIR; J01403; J01403;
DR PDB; 1SFP; 24-JUN-98.
DR InterPro; IPR000124; -
DR InterPro; IPR000859; -
DR Pfam; PF00431; CUB; 1.
DR PROSITE; PS00985; SPERMADHESIN_1; 1.
DR PROSITE; PS00986; SPERMADHESIN_2; 1.
DR PROSITE; PS01180; CUB; 1.
DR Growth factor; Signal; 3D-structure.
KW SIGNAL
FT SIGNAL 1 20
FT CHAIN 21 134 ACIDIC SEMINAL FLUID PROTEIN.
FT DOMAIN 30 131
FT DISULFID 30 51
FT DISULFID 74 95
FT CONFLICT 43 43 T -> H (IN REF. 2).
SQ SEQUENCE 134 AA; 15036 MW; 339BCFF8637D64C0 CRC64;

Query Match      10.7%; Score 6; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 VKYIRE 55
Db 111 VKYIRE 116
```

```

RT  "The proteinase yscA-inhibitor, IA3, gene. Studies of cytoplasmic
RL  proteinase inhibitor deficiency on yeast physiology.";
RN  FEBS Lett. 283:78-84(1991).
RP  [2]
RC  SEQUENCE FROM N.A.
RA  STRAIN-S288C / AB972;
RA  Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL  Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN  [3]
RP  SEQUENCE.
RA  Biedermann K., Montali U., Martin B., Svendsen I., Ottesen M.;
RT  "The amino acid sequence of proteinase A inhibitor 3 from baker's
RL  yeast.";
RL  Carlsberg Res. Commun. 45:225-235(1980).
CC  -!- FUNCTION: INHIBITOR FOR YEAST PROTEASE A (YSCA).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X60050; CAA42650.1; -.
DR  EMBL; Z49808; CAA89507.1; -.
DR  PIR; A01334; IABY3
DR  PIR; S16692; S16692.
DR  SGD; S0004786; PAI3.
KW  Protease inhibitor; Acetylation.
FT  MOD_RES 1 1 ACETYLATION.
FT  DOMAIN 2 23 NEARLY ALL OF THE INHIBITORY ACTIVITY IS
FT  PRESENT IN THIS REGION.
SQ  SEQUENCE 68 AA; 7707 MW; 8295BF7AF3E4A01A CRC64;

Query Match 10.7%; Score 6; DB 1; Length 68;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSKKEL 13
Db 14 SSKKEL 19

RESULT 12
RS18_UREPA STANDARD; PRT; 79 AA.
AC Q9PPT6;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 30S RIBOSOMAL PROTEIN S18.
GN RPSR OR RPS18 OR U0552.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SEROVAR 3;
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
urealyticum.";
RL Nature 407:757-762(2000).
CC -!- FUNCTION: THIS PROTEIN HAS BEEN IMPLICATED IN AMINOACYL-TRANSFER
CC RNA BINDING. IT APPEARS TO BE SITUATED AT THE DECODING SITE OF
CC MESSENGER RNA (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; AE002154; AAF30965.1; -.
DR  InterPro; IPR001648; -.
DR  Pfam; PF01084; Ribosomal_S18; 1.
DR  PRINTS; PR00974; RIBOSOMALS18
DR  PROSITE; PS00057; RIBOSOMAL_S18; FALSE_NEG.
KW  Ribosomal protein; RNA-binding.
SQ  SEQUENCE 79 AA; 9136 MW; E4697FD3F03F5AC4 CRC64;

Query Match 10.7%; Score 6; DB 1; Length 79;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 LLPYVK 33
Db 73 LLPYVK 78

RESULT 13
HFQ_HAEIN STANDARD; PRT; 90 AA.
AC P44437;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HOST FACTOR-I PROTEIN (HF-I).
GN HFQ OR HI0411.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: REQUIRED FOR PHAGE Q BETA RNA-DIRECTED SYNTHESIS OF
CC COMPLEMENTARY MINUS-STRAND RNA (BY SIMILARITY).
CC -!- SUBUNIT: HOMOHXAMER (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; U32724; AAC22070.1; -.
DR  TIGR; HI0411; -.
KW  INIT_MET 0 0 BY SIMILARITY.
FT  INIT_MET 0 0 BY SIMILARITY.
SQ  SEQUENCE 90 AA; 10118 MW; IC91DAECF1F7F029 CRC64;

```

SUBFAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X64603; CAA43887.1; -  
PIR; A44337; A44337.  
HSSP; P17119; 3KAR.  
InterPro; IPR001752; -  
Pfam; PF00225; kinesin; 1.  
PRINTS; PR00380; KINESINHEAVY.  
PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
PROSITE; PS0067; KINESIN\_MOTOR\_DOMAIN2; 1.  
KW Motor protein; Microtubules; ATP-binding; Coiled coil.  
FT DOMAIN 175 425 COILED COIL (POTENTIAL).  
FT DOMAIN 426 770 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).  
FT NP\_BIND 514 521 ATP (BY SIMILARITY).  
SQ SEQUENCE 770 AA; 85800 MW; EFD0F0FF39B5C7EB CRC64;

Query Match 12.5%; Score 7; DB 1; Length 770;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 KEKLRR 16  
| | | | | | | |  
Db 398 KEKLRR 404

RESULT 9  
YMC9\_YEAST  
ID YMC9\_YEAST STANDARD; PRT; 838 AA.  
AC Q03714;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE HYPOTHETICAL 96.7 KDA PROTEIN IN NDC1-TSAL INTERGENIC REGION.  
GN YML029W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; Z46659; CAA86626.1; -  
SGD; S000491; YML029W.  
InterPro; IPR000626; -  
PROSITE; PS00553; UBIQUITIN\_2; UNKNOWN\_1.  
KW Hypothetical protein; Transmembrane.  
FT POTNSMEM 224 244 POTENTIAL.  
FT DOMAIN 259 318 UBIQUITIN-LIKE.  
FT TRANSMEM 532 552 POTENTIAL.  
FT TRANSMEM 553 573 POTENTIAL.  
FT TRANSMEM 764 784 POTENTIAL.

SQ SEQUENCE 838 AA; 96653 MW; 9B93ECA6C5421FD6 CRC64;

Query Match 12.5%; Score 7; DB 1; Length 838;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 27 TLLPYVK 33  
| | | | | | | |  
Db 778 TLLPYVK 784

RESULT 10  
VG1A\_BPPZA  
ID VG1A\_BPPZA STANDARD; PRT; 59 AA.  
AC P06947;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 01-JAN-1988 (Rel. 06, Last annotation update)  
DE EARLY PROTEIN GP1A.  
GN 1A.  
OS Bacteriophage PZA.  
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Podoviridae.  
OX NCBI\_TaxID=10757;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86056991; PubMed=3934048;  
RA Paces V., Vilek C., Urbanek P., Hostomsky Z.;  
RT "Nucleotide sequence of the major early region of Bacillus subtilis phage PZA, a close relative of phi 29.";  
RL Gene 38:45-56(1985).  
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; M11813; AAA88481.1; -  
PIR; A24528; ERBP1A.  
KW Early protein.  
SQ SEQUENCE 59 AA; 6865 MW; FCC525137B72D831 CRC64;

Query Match 10.7%; Score 6; DB 1; Length 59;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 38 DAASVL 43  
| | | | | | | |  
Db 28 DAASVL 33

RESULT 11  
IPA3\_YEAST  
ID IPA3\_YEAST STANDARD; PRT; 68 AA.  
AC P01094;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE PROTEASE A INHIBITOR 3 (PROTEINASE INHIBITOR 1(A)3).  
GN PAI3 OR YMR174C OR YMR010.04C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=91243884; PubMed=2037077;  
RA Schu P., Wolf D.H.;

GN POTA OR MG042.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 33530 / G-37; PubMed=7569993;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.F., Dougherty B.A., Bock K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium."  
 RL Science 270:397-403(1995).  
 CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
 CC FOR SPERMIDINE/PUTRESCINE. PROBABLY RESPONSIBLE FOR ENERGY  
 CC COUPLING TO THE TRANSPORT SYSTEM (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: U39684; AAC71258.1; -;  
 DR HSSP: P13569; INBD.  
 DR TIGR: MG042; -;  
 DR InterPro: IPR001617; -;  
 DR Pfam: PF00005; ABC\_tran; 1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 KW Transport; ATP-binding; Membrane.  
 FT NP\_BIND 40 47 ATP (POTENTIAL).  
 SQ SEQUENCE 559 AA; 65192 MW; DB8F04B26A110F6F CRC64;  
 -----  
 Query Match 12.5%; Score 7; DB 1; Length 559;  
 Best Local Similarity 100.0%; Pred. No. 9;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 10 KEKLRRE 16  
 |  
 Db 173 KEKLRRE 179  
 |  
 RESULT 7  
 CN1A\_MOUSE STANDARD; PRT; 565 AA.  
 ID CN1A\_MOUSE  
 AC Q61481; Q35388;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CALCIUM/CALMODULIN-DEPENDENT 3', 5'-CYCLIC NUCLEOTIDE PHOSPHODIESTERASE  
 DE 1A (EC 3.1.4.17) (CAM-PDE 1A) (61 KDA CAM-PDE).  
 GN PDE1A.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (PDE1A2).  
 RC STRAIN-BALB/C; TISSUE-Brain;  
 RA Yan C., Sonnenburg W.K., Zhao A.Z., Kwak K.S., Beavo J.A.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-262 FROM N.A. (PDE1A1).

RC TISSUE-Heart;  
 RA Sonnenburg W.K., Rybalkin S.D., Bornfeldt K.E., Kwak K.S.,  
 RA Rybalkina I., Beavo J.A.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: HAS A HIGHER AFFINITY FOR CGMP THAN FOR CAMP.  
 CC -1- CATALYTIC ACTIVITY: GUANOSINE 3', 5'-CYCLIC PHOSPHATE + H(2)O =  
 CC GUANOSINE 5'-PHOSPHATE.  
 CC -1- ENZYME REGULATION: TYPE I PDE ARE ACTIVATED BY THE BINDING OF  
 CC CALMODULIN IN THE PRESENCE OF CA(2+).  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: THERE ARE DIFFERENT ISOFORMS PRODUCED BY  
 CC ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ISOFORM  
 CC PDE1A2.  
 CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: U56649; AAB03319.1; -;  
 DR EMBL: AF023529; AAB81952.1; -;  
 DR MGD: MGI:1201792; Pde1a.  
 DR InterPro: IPR002073; -;  
 DR Pfam: PF00233; PDEase; 1.  
 DR PRINTS: PR00387; PDIESTERASE1.  
 DR PROSITE: PS00126; PDEASE\_1; 1.  
 KW Hydrolyase; CAMP; cGMP; Multigene family; Alternative splicing;  
 KW Calmodulin-binding.  
 FT DOMAIN 44 64 CALMODULIN-BINDING.  
 FT DOMAIN 213 535 CATALYTIC (BY SIMILARITY).  
 FT VARSPPLIC 1 54 MVGSSTSSHWIAPVRNINMGSTDDIEELEENATYKYLIG  
 FT EQTEKMQORLKI -> MDEYVTRKKHLQRPPIR (IN  
 FT ISOFORM PDE1A1).  
 SQ SEQUENCE 565 AA; 64470 MW; 66A0749774967FE6 CRC64;  
 -----  
 Query Match 12.5%; Score 7; DB 1; Length 565;  
 Best Local Similarity 100.0%; Pred. No. 9.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 39 AASVLEA 45  
 |  
 Db 79 AASVLEA 85  
 |  
 RESULT 8  
 KLPA\_EMENI STANDARD; PRT; 770 AA.  
 ID KLPA\_EMENI  
 AC P28739;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE KINESIN-LIKE PROTEIN KLPA.  
 GN KLPA.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; Emericella.  
 OX NCBI\_TaxID=5072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GB20;  
 RX MEDLINE=93107178; PubMed=8416986;  
 RA O'Connell M.J., Meluh P.B., Rose M.D., Morris N.R.;  
 RT "Suppression of the bimC4 mitotic spindle defect by deletion of klpa,  
 RT a gene encoding a KAR3-related kinesin-like protein in Aspergillus  
 RT nidulans."  
 RL J. Cell Biol. 120:153-162(1993).  
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD

```

RX MEDLINE=93107074; PubMed=7678006;
RA Sonnenburg W.K., Seger D., Beavo J.A.;
RT "Molecular cloning of a cDNA encoding the '61-kDa'
RT calmodulin-stimulated cyclic nucleotide phosphodiesterase."
RT Tissue-specific expression of structurally related isoforms."
RL J. Biol. Chem. 268:645-652(1993).
RN [2]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91329365; PubMed=1651111;
RA Charbonneau H., Kumar S., Novack J.P., Blumenthal D.K., Griffin P.R.,
RA Shabanowitz J., Hunt D.F., Beavo J.A., Walsh K.A.;
RT "Evidence for domain organization within the 61-kDa
RT calmodulin-dependent cyclic nucleotide phosphodiesterase from bovine
RT brain."
RL Biochemistry 30:7931-7940(1991).
RN [3]
RP SEQUENCE OF 193-426.
RC TISSUE=Brain;
RX MEDLINE=87092242; PubMed=3025833;
RA Charbonneau H., Beier N., Walsh K.A., Beavo J.A.;
RT "Identification of a conserved domain among cyclic nucleotide
RT phosphodiesterases from diverse species."
RL Proc. Natl. Acad. Sci. U.S.A. 83:9308-9312(1986).
CC -!- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC GUANOSINE 5'-PHOSPHATE.
CC -!- ENZYME REGULATION: TYPE I PDE ARE ACTIVATED BY THE BINDING OF
CC CALMODULIN IN THE PRESENCE OF CA(2+).
CC -!- SUBUNIT: HOMODIMER.
CC -!- ALTERNATIVE PRODUCTS: THERE ARE PROBABLY DIFFERENT ISOFORMS
CC PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF
CC ISOFORM PDE1A2.
CC -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M90358; AAA74560.1; -
DR PIR; A26650; A26650.
DR PIR; A40282; A40282.
DR PIR; A45334; A45334.
DR InterPro; IPR002073; -.
DR Pfam; PF00233; PDEase; 1.
DR PRINTS; PR00387; PDIESTERASE1.
DR PROSITE; PS00126; PDEASE_I; 1.
DR PROSITE; PS00126; PDEASE_I; 1.
KW Hydrolase; cAMP; cGMP; Multigene family; Alternative splicing;
KW Calmodulin-binding.
FT INIT_MET 0 0
FT DOMAIN 23 43 CALMODULIN-BINDING.
FT DOMAIN 192 500 CATALYTIC (BY SIMILARITY).
FT CONFLICT 236 236 H -> G (IN REF. 3).
FT CONFLICT 320 320 N -> W (IN REF. 3).
SQ SEQUENCE 529 AA; 60712 MW; 24DB74C22A2AE06F CRC64;

Query Match 12.5%; Score 7; DB 1; Length 529;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 39 AASVLEA 45
Db 58 AASVLEA 64
|||||
RESULT 5
CNIA_HUMAN
ID CNIA_HUMAN STANDARD; PRT; 534 AA.

RX MEDLINE=93107074; PubMed=7678006;
RA Sonnenburg W.K., Seger D., Beavo J.A.;
RT "Molecular cloning of a cDNA encoding the '61-kDa'
RT calmodulin-stimulated cyclic nucleotide phosphodiesterase."
RT Tissue-specific expression of structurally related isoforms."
RL J. Biol. Chem. 268:645-652(1993).
RN [2]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91329365; PubMed=1651111;
RA Charbonneau H., Kumar S., Novack J.P., Blumenthal D.K., Griffin P.R.,
RA Shabanowitz J., Hunt D.F., Beavo J.A., Walsh K.A.;
RT "Evidence for domain organization within the 61-kDa
RT calmodulin-dependent cyclic nucleotide phosphodiesterase from bovine
RT brain."
RL Biochemistry 30:7931-7940(1991).
RN [3]
RP SEQUENCE OF 193-426.
RC TISSUE=Brain;
RX MEDLINE=87092242; PubMed=3025833;
RA Charbonneau H., Beier N., Walsh K.A., Beavo J.A.;
RT "Identification of a conserved domain among cyclic nucleotide
RT phosphodiesterases from diverse species."
RL Proc. Natl. Acad. Sci. U.S.A. 83:9308-9312(1986).
CC -!- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC GUANOSINE 5'-PHOSPHATE.
CC -!- ENZYME REGULATION: TYPE I PDE ARE ACTIVATED BY THE BINDING OF
CC CALMODULIN IN THE PRESENCE OF CA(2+).
CC -!- SUBUNIT: HOMODIMER.
CC -!- ALTERNATIVE PRODUCTS: THERE ARE PROBABLY DIFFERENT ISOFORMS
CC PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF
CC ISOFORM PDE1A2.
CC -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M90358; AAA74560.1; -
DR PIR; A26650; A26650.
DR PIR; A40282; A40282.
DR PIR; A45334; A45334.
DR InterPro; IPR002073; -.
DR Pfam; PF00233; PDEase; 1.
DR PRINTS; PR00387; PDIESTERASE1.
DR PROSITE; PS00126; PDEASE_I; 1.
KW Hydrolase; cAMP; cGMP; Multigene family; Alternative splicing;
KW Calmodulin-binding.
FT INIT_MET 0 0
FT DOMAIN 23 43 CALMODULIN-BINDING.
FT DOMAIN 192 500 CATALYTIC (BY SIMILARITY).
FT CONFLICT 236 236 H -> G (IN REF. 3).
FT CONFLICT 320 320 N -> W (IN REF. 3).
SQ SEQUENCE 529 AA; 60712 MW; 24DB74C22A2AE06F CRC64;

Query Match 12.5%; Score 7; DB 1; Length 529;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 39 AASVLEA 45
Db 58 AASVLEA 64
|||||
RESULT 5
CNIA_HUMAN
ID CNIA_HUMAN STANDARD; PRT; 534 AA.
```

```

AC P54750;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALCIUM/CALMODULIN-DEPENDENT 3',5'-CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
DE 1A (EC 3.1.4.17) (CAM-PDE 1A) (61 KDA CAM-PDE) (HCAM-1).
GN PDE1A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96132810; PubMed=8557689;
RA Loughney K., Martins T.J., Harris E.A.S., Sadhu K., Hicks J.B.,
RA Sonnenburg W.K., Beavo J.A., Ferguson K.;
RT "Isolation and characterization of cDNAs corresponding to two human
RT calcium, calmodulin-regulated, 3',5'-cyclic nucleotide
RT phosphodiesterases."
RL J. Biol. Chem. 271:796-806(1996).
CC -!- FUNCTION: HAS A HIGHER AFFINITY FOR CGMP THAN FOR CAMP.
CC -!- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC GUANOSINE 5'-PHOSPHATE.
CC -!- ENZYME REGULATION: TYPE I PDE ARE ACTIVATED BY THE BINDING OF
CC CALMODULIN IN THE PRESENCE OF CA(2+).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- ALTERNATIVE PRODUCTS: THERE ARE PROBABLY DIFFERENT ISOFORMS
CC PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF
CC ISOFORM PDE1A3.
CC -!- TISSUE SPECIFICITY: SEVERAL TISSUES, INCLUDING BRAIN, KIDNEY,
CC TESTES, AND HEART.
CC -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U40370; AAC50436.1; -
DR MIM; 171890; -.
DR InterPro; IPR002073; -.
DR Pfam; PF00233; PDEase; 1.
DR PRINTS; PR00387; PDIESTERASE1.
DR PROSITE; PS00126; PDEASE_I; 1.
KW Hydrolase; cAMP; cGMP; Multigene family; Alternative splicing;
KW Calmodulin-binding.
FT INIT_MET 0 0 BY SIMILARITY.
FT DOMAIN 23 43 CALMODULIN-BINDING (BY SIMILARITY).
FT DOMAIN 192 514 CATALYTIC (BY SIMILARITY).
SQ SEQUENCE 534 AA; 61120 MW; 8398FC451608A720 CRC64;

Query Match 12.5%; Score 7; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 39 AASVLEA 45
Db 58 AASVLEA 64
|||||
RESULT 6
POTA_MYCGE
ID POTA_MYCGE STANDARD; PRT; 559 AA.
AC P47288;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SPERMIDINE/PUTRESCINE TRANSPORT ATP-BINDING PROTEIN POTA HOMOLOG.
```

DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HYPOTHETICAL PROTEIN PA4543.  
 GN PA4543.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 RN [2]  
 RP SEQUENCE OF 198-242 FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=93225810; PubMed=8097014;  
 RA Hobbs M., Collie E.S., Free P.D., Livingston S.P., Mattick J.S.;  
 RT "Pils and PilR, a two-component transcriptional regulatory system  
 RT controlling expression of type 4 fimbriae in Pseudomonas  
 RT aeruginosa.";  
 RL Mol. Microbiol. 7:669-682(1993).  
 CC -1- SIMILARITY: BELONGS TO THE UPF0124 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE004868; AAG07931.1; -  
 DR EMBL; L06013; AAA87639.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 242 AA; 26047 MW; AD4470FF676FF51D CRC64;

Query Match 12.5%; Score 7; DB 1; Length 242;  
 Best Local Similarity 100.0%; Pred. No. 4.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VLEATVD 48  
 Db 130 VLEATVD 136  
 |||||

RESULT 3  
 DPFF\_HAEIN  
 ID DPFF\_HAEIN STANDARD; PRT; 327 AA.  
 AC P45094;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE DIPEPTIDE TRANSPORT ATP-BINDING PROTEIN DPFF.  
 GN DPFF OR H11184.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Buit C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uutterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 RT influenzae Rd.";  
 RL Science 269:496-512(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20 / ATCC 51907;  
 RX MEDLINE=96134971; PubMed=8550458;  
 RA Preston A., Maskell D., Johnson A., Moxon E.R.;  
 RT "Altered lipopolysaccharide characteristic of the I69 phenotype in  
 RT Haemophilus influenzae results from mutations in a novel gene, isn.";  
 RL J. Bacteriol. 178:396-402(1996).  
 CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
 CC FOR DIPEPTIDES. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO  
 CC THE TRANSPORT SYSTEM (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U32798; AAC22837.1; -  
 DR EMBL; U17295; AAA95975.1; -  
 DR TIGR; H11184; -  
 DR InterPro; IPR001617; -  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 DR Peptide transport; Transport; Inner membrane; ATP-binding.  
 FT NP\_BIND 54 61 ATP (POTENTIAL)  
 SQ SEQUENCE 327 AA; 36917 MW; 0BB0BBDE197DA9BE CRC64;

Query Match 12.5%; Score 7; DB 1; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 5.7;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRRERIK 19  
 Db 271 LRRERIK 277  
 |||||

RESULT 4  
 CNIA\_BOVIN  
 ID CNIA\_BOVIN STANDARD; PRT; 529 AA.  
 AC P14100;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CALCIUM/CALMODULIN-DEPENDENT 3', 5'-CYCLIC NUCLEOTIDE PHOSPHODIESTERASE  
 DE 1A (EC 3.1.4.17) (CAM-PDE 1A) (61 KDA CAM-PDE).  
 GN PDE1A.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:50:38 ; Search time 23.18 Seconds  
(without alignments)  
82.757 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_189  
Perfect score: 56  
Sequence: 1 KKSLLHSSKEKLREIRKY.....NDAAASVLEATVDYVKYIREK 56

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	7	12.5	169	1 Y769_METJA	Q58179 methanococ
2	7	12.5	242	1 Y9E3_PSEAE	P33663 pseudomonas
3	7	12.5	327	1 DPPE_HAEIN	P45094 haemophilus
4	7	12.5	529	1 CN1A_BOVIN	P14100 bos taurus
5	7	12.5	534	1 CN1A_HUMAN	P54750 homo sapien
6	7	12.5	559	1 POTA_MYCGE	P47288 mycoplasma
7	7	12.5	565	1 CN1A_MOUSE	Q61481 mus musculus
8	7	12.5	770	1 KLP4_EMENI	P28739 emericella
9	7	12.5	838	1 YMC9_YEAST	Q03714 saccharomyc
10	6	10.7	59	1 VGI1A_BPZA	P06947 bacterioph
11	6	10.7	68	1 IPA3_YEAST	P01094 saccharomyc
12	6	10.7	79	1 RS18_UREPA	Q9PPT8 ureaplasma
13	6	10.7	90	1 HFQ_HAEIN	P44437 haemophilus
14	6	10.7	93	1 YB4C_SCHPO	O14358 schizosacch
15	6	10.7	134	1 ASFP_BOVIN	P29392 bos taurus
16	6	10.7	145	1 RR17_ORYZA	Q9ZST1 oryza sativ
17	6	10.7	145	1 Y004_TREPA	O83050 treponema p
18	6	10.7	150	1 LC4S_HUMAN	Q16873 homo sapien
19	6	10.7	153	1 MAL_CANFA	Q28296 canis fami
20	6	10.7	153	1 MAL_MOUSE	P21145 homo sapien
21	6	10.7	153	1 MAL_MOUSE	O09198 mus musculus
22	6	10.7	153	1 MAL_RAT	O64349 rattus norv
23	6	10.7	159	1 GRE4_BUCAI	P57464 buchnera ap
24	6	10.7	160	1 YMT0_YEAST	Q04210 saccharomyc
25	6	10.7	163	1 PGPA_HAEIN	P44157 haemophilus
26	6	10.7	177	1 PUR6_PYRHO	O58058 pyrococcus
27	6	10.7	177	1 NOHB_ECOLI	P31062 escherichia
28	6	10.7	181	1 TERS_LAMBD	P03707 bacterioph
29	6	10.7	189	1 NOHA_ECOLI	P31061 escherichia
30	6	10.7	236	1 PLC1_BOVIN	P09611 bos taurus
31	6	10.7	236	1 PLC_SHEEP	P16038 ovis aries
32	6	10.7	239	1 RP35_BACTK	P26763 bacillus th
33	6	10.7	242	1 YTXE_BACSU	P39064 bacillus su

34	6	10.7	248	1 DSBG_ECOLI	P77202 escherichia
35	6	10.7	250	1 VGLL_HSV6U	P52508 human herpe
36	6	10.7	250	1 VGLL_HSV6Z	P52526 human herpe
37	6	10.7	258	1 VGLL_GPCMV	O92277 guinea pig
38	6	10.7	268	1 RPNA_YEAST	P38886 saccharomyc
39	6	10.7	268	1 VANY_ENTFA	Q47746 enterococcu
40	6	10.7	274	1 OSA4_BORBU	O4851 borrelia bu
41	6	10.7	274	1 YA99_SCHPO	O09787 schizosacch
42	6	10.7	289	1 IPYR_BOVIN	P37980 bos taurus
43	6	10.7	289	1 IPYR_HUMAN	Q15181 homo sapien
44	6	10.7	291	1 YXJO_BACSU	P55181 bacillus su
45	6	10.7	292	1 YN19_MYCTU	P71893 mycobacteri

ALIGNMENTS

RESULT 1					
Y769_METJA					
ID Y769_METJA	STANDARD;	PRT;	169 AA.		
AC Q58179;					
DT 01-NOV-1997 (Rel. 35, Created)					
DT 01-NOV-1997 (Rel. 35, Last sequence update)					
DT 01-NOV-1997 (Rel. 35, Last annotation update)					
DE HYPOTHETICAL PROTEIN MJ0769.					
GN MJ0769.					
OS Methanococcus jannaschii.					
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;					
OC Methanococcus.					
OX NCBI_TaxID=2190;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;					
RX MEDLINE=96337999; PubMed=8688087;					
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,					
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Adams M.D., Reich C.I.,					
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Merrick J.M., Glodok A.,					
RA Overbeek R., Kirkness E.F., Weinstock K.G., Weinstock J.M., Nguyen D.,					
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Sadow P.W., Hanna M.C.,					
RA Uterback T.R., Kelley J.M., Peterson J.D., Hurst M.A., Kaine B.P., Borodovsky M.,					
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;					
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus					
RL Science 273:1058-1073(1996).					
CC -----					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC the European Bioinformatics Institute. There are no restrictions on its					
CC use by non-profit institutions as long as its content is in no way					
CC modified and this statement is not removed. Usage by and for commercial					
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC or send an email to license@isb-sib.ch).					
CC -----					
DR EMBL; U67522; AAB98774.1; -					
DR TIGR; MJ0769; -					
KW Hypothetical protein.					
SQ SEQUENCE 169 AA; 20167 MW; CIDEFDBIEFI23898 CRC64;					

Query Match 12.5%; Score 7; DB 1; Length 169;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	43	LEATVDY 49
Db	144	LEATVDY 150

RESULT 2					
Y9E3_PSEAE					
ID Y9E3_PSEAE	STANDARD;	PRT;	242 AA.		
AC P33663;					



**THIS PAGE BLANK (USPTO)**



A:Residues: 1-390 <STO>  
 A:Cross-references: GB:AE004885; GB:AE004091; NID:g9950968; PIDN:AAG08108.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA4722  
 C:Superfamily: aspartate transaminase

Query Match 12.5%; Score 7; DB 2; Length 390;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 DAASVLE 44  
 |||||  
 Db 215 DAASVLE 221

RESULT 12  
 T48593  
 hypothetical protein T22N19.110 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: T48593  
 R:Byan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, submitted to the Protein Sequence Database, April 2000  
 A:Reference number: 224490  
 A:Accession: T48593  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-443 <BEV>  
 A:Cross-references: EMBL:AL163572  
 A:Experimental source: cultivar Columbia; BAC clone T22N19  
 C:Genetics:  
 A:Map position: 5  
 A:Introns: 23/3; 125/3; 196/3; 240/3  
 A:Note: T22N19.110

Query Match 12.5%; Score 7; DB 2; Length 443;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LRRERIK 19  
 |||||  
 Db 227 LRRERIK 233

RESULT 13  
 P64204  
 spermidine/putrescine transport ATP-binding protein potA homolog - Mycoplasma genitalium  
 C:Species: Mycoplasma genitalium  
 C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 17-Mar-2000  
 C:Accession: P64204  
 R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.; M.; Fuhrman, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J., C.A.; Venter, J.C.  
 Science 270, 397-403, 1995  
 A:Title: The minimal gene complement of Mycoplasma genitalium.  
 A:Reference number: A64200; MUID:96026346  
 A:Accession: P64204  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-486 <TIGR>  
 A:Cross-references: GB:U39683; GB:I43967; NID:g1045711; PID:g1045714; TIGR:MG042  
 A:Experimental source: strain G-37  
 C:Genetics:

A:Genetic code: SGC3  
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
 C:Keywords: ATP  
 F:1-351/Domain: ATP-binding cassette homology #status atypical <ABCL>

Query Match 12.5%; Score 7; DB 2; Length 486;

Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KEKLRR 16  
 |||||  
 Db 100 KEKLRR 106

RESULT 14  
 A40283  
 3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17), calmodulin-dependent, 59K ca  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 21-Jan-2000  
 C:Accession: A40283  
 R:Novack, J.P.; Charbonneau, H.; Bentley, J.K.; Walsh, K.A.; Beavo, J.A.  
 Biochemistry 30, 7940-7947, 1991  
 A:Title: Sequence comparison of the 63-, 61-, and 59-kDa calmodulin-dependent cyclic  
 A:Reference number: A40283; MUID:91329366  
 A:Accession: A40283  
 A:Molecule type: protein  
 A:Residues: 1-491 <NOV>  
 C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase 1B, calmodulin-dependent; 3'  
 C:Keywords: calmodulin binding; cardiac muscle; heart; phosphoric diester hydrolase  
 F:196-396/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNP>

Query Match 12.5%; Score 7; DB 2; Length 491;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVLEA 45  
 |||||  
 Db 43 AASVLEA 49

RESULT 15  
 T14783  
 hypothetical protein DKFZp586G0221.1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jan-2000  
 C:Accession: T14783  
 R:Ottenwälder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, August 1999  
 A:Reference number: Z18184  
 A:Accession: T14783  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-519 <OFT>  
 A:Cross-references: EMBL:AL110263  
 A:Experimental source: adult uterus; clone DKFZp586G0221  
 C:Genetics:  
 A:Note: DKFZp586G0221.1  
 C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase 1B, calmodulin-dependent; 3'  
 F:202-419/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNP>

Query Match 12.5%; Score 7; DB 2; Length 519;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVLEA 45  
 |||||  
 Db 43 AASVLEA 49

Search completed: September 15, 2001, 12:48:52  
 Job time: 222 sec

C:Accession: H85067  
 R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999  
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
 A:Reference number: A85001; MUID:20083488  
 A:Accession: H85067  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-250 <STO>  
 A:Cross-references: GB:NC\_001268; NID:g7267300; PIDN:CAB81082.1; GSPDB:GN00140  
 C:Genetics:  
 A:Gene: AT4g05400  
 A:Map position: 4  
 C:Superfamily: Arabidopsis thaliana hypothetical protein P7J7.80

Query Match 12.5%; Score 7; DB 2; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKELR 14  
 DB 244 SSKELR 250  
 |||||

RESULT 8  
 B71312  
 probable GTP-binding protein (era) - syphilis spirochete  
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
 C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 02-Feb-2001  
 C:Accession: B71312  
 R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDowell, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
 Science 281, 375-388, 1998  
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
 A:Reference number: A71250; MUID:98332770  
 A:Accession: B71312  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-319 <COL>  
 A:Cross-references: GB:AE001229; GB:AE000520; NID:g3322827; PIDN:AAC65525.1; PID:g332283  
 A:Experimental source: strain Nichols  
 C:Genetics:  
 A:Gene: TP0541  
 C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
 C:Keywords: GTP binding; nucleotide binding; P-loop  
 F:11-130/Domain: translation elongation factor Tu homology <ETU>  
 F:17-24/Region: nucleotide-binding motif A (P-loop)  
 F:127-130/Region: GTP-binding NKXD motif  
 F:176-178/Region: GTP-binding SAK/L motif

Query Match 12.5%; Score 7; DB 2; Length 319;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LRTLLPY 31  
 DB 287 LRTLLPY 293  
 |||||

RESULT 9  
 H72472  
 hypothetical protein APE2425 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: H72472  
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahashi, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kawauchi, S.; Lory, S.; Olson, M.V.  
 DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1.  
 A:Reference number: A72450; MUID:99310339

A:Accession: H72472  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-326 <KAW>  
 A:Cross-references: DBJ:AP000064; NID:g5105945; PIDN:BA081440.1; PID:g5106129  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE2425  
 C:Superfamily: conserved hypothetical protein MJ1427

Query Match 12.5%; Score 7; DB 2; Length 326;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AASVLEA 45  
 DB 292 AASVLEA 298  
 |||||

RESULT 10  
 E64188  
 dipeptide transport ATP-binding protein dpfF - Haemophilus influenzae (strain Rd KW20)  
 C:Species: Haemophilus influenzae  
 C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 20-Aug-1999  
 C:Accession: E64188  
 R:Flischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geohagen, N.S.M. Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.C.  
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; MUID:95350630  
 A:Accession: E64188  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-327 <TIGR>  
 A:Cross-references: GB:U32798; GB:L42023; NID:gl574110; PIDN:AAC22837.1; PID:gl574111  
 C:Genetics:  
 A:Gene: dpfF  
 C:Function:  
 A:Description: probably responsible for energy-coupling to the transport system  
 C:Superfamily: inner membrane protein malk; ATP-binding cassette homology  
 C:Keywords: ATP; binding protein-dependent transport system; membrane protein; nucleoside binding; ATP-binding cassette homology <ABC>  
 F:37-237/Domain: ATP-binding cassette homology <ABC>  
 F:54-62/Region: nucleotide-binding motif A (P-loop)  
 F:181-185/Region: nucleotide-binding motif B

Query Match 12.5%; Score 7; DB 2; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LRRERIK 19  
 DB 271 LRRERIK 277  
 |||||

RESULT 11  
 D83057  
 probable aminotransferase PA4722 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: D83057  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: D83057  
 A:Status: preliminary  
 A:Molecule type: DNA

C:Function:  
 A:Description: confers resistance against Fusarium oxysporum  
 C:Superfamily: disease resistance protein RPS2; leucine-rich alpha-2-glycoprotein repeat

Query Match 14.3%; Score 8; DB 2; Length 1220;  
 Best Local Similarity 100.0%; Pred. No. 3.5; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0;

QY 23 EQLRTLTP 30  
 |||||  
 DB 554 EQLRTLTP 561

RESULT 3  
 T06404  
 resistance complex protein I2C-2 - tomato  
 C:Species: Lycopersicon esculentum (tomato)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 17-Nov-2000  
 C:Accession: T06404  
 R:Ort, N.; Eshed, Y.; Paran, I.; Presting, G.; Aviv, D.; Tanksley, S.; Zamir, D.; Fluhr,  
 Plant Cell 9, 521-532, 1997  
 A:Title: The I2C family from the wilt disease resistance locus I2 belongs to the nucleot  
 A:Reference number: Z15652; MUID:97290204  
 A:Accession: T06404  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1240 <ORI>  
 A:Cross-references: EMBL:AF004879; NID:g2258316; PIDN:AAB63275.1; PID:g2258317  
 C:Genetics:  
 A:Gene: I2C-2  
 A:Map position: 11  
 C:Function:  
 A:Description: confers resistance against Fusarium oxysporum  
 C:Superfamily: disease resistance protein RPS2; leucine-rich alpha-2-glycoprotein repeat

Query Match 14.3%; Score 8; DB 2; Length 1240;  
 Best Local Similarity 100.0%; Pred. No. 3.6; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0;

QY 23 EQLRTLTP 30  
 |||||  
 DB 545 EQLRTLTP 552

RESULT 4  
 A64396  
 hypothetical protein MJ0769 - Methanococcus jannaschii  
 C:Species: Methanococcus jannaschii  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 C:Accession: A64396  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
 Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
 A:Reference number: A64300; MUID:96337999  
 A:Accession: A64396  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-169 <BOL>  
 A:Cross-references: GB:U67522; GB:L77117; NID:g2826315; PIDN:AAB98774.1; PID:g1499589; T  
 C:Genetics:  
 A:Map position: REV690989-690480

Query Match 12.5%; Score 7; DB 2; Length 169;  
 Best Local Similarity 100.0%; Pred. No. 7.3; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0;

QY 43 LEATVDY 49

Db 144 LEATVDY 150  
 |||||  
 RESULT 5  
 A70527  
 hypothetical protein Rv0328 - Mycobacterium tuberculosis (strain H37RV)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: A70527  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
 A:Reference number: A70500; MUID:98295987  
 A:Accession: A70527  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-200 <COL>  
 A:Cross-references: GB:Z96800; GB:AL123456; NID:g3261800; PIDN:CAB09609.1; PID:g21939  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: RV0328

Query Match 12.5%; Score 7; DB 2; Length 200;  
 Best Local Similarity 100.0%; Pred. No. 8.4; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0;

QY 42 VLEATVD 48  
 |||||  
 DB 88 VLEATVD 94

RESULT 6  
 EB3077  
 conserved hypothetical protein PA4543 [imported] - Pseudomonas aeruginosa (strain PAO  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: EB3077  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: EB3077  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-242 <STO>  
 A:Cross-references: GB:AE004868; GB:AE004091; NID:g9950785; PIDN:ANG07931.1; GSPDB:GN  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA4543  
 C:Superfamily: Streptomyces coelicolor hypothetical protein SC4A10.14c

Query Match 12.5%; Score 7; DB 2; Length 242;  
 Best Local Similarity 100.0%; Pred. No. 9.9; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 0;

QY 42 VLEATVD 48  
 |||||  
 DB 130 VLEATVD 136

RESULT 7  
 H85067  
 hypothetical protein AT4g05400 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: September 15, 2001, 12:48:49 ; Search time 45.39 Seconds  
(without alignments)  
93.981 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_189  
Perfect score: 56  
Sequence: 1 KKISLLHSSKEKLRRRIKY.....NDAASVLEATVDYVKVIREK 56

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	14.3	703	2 G75638	nodulation protein
2	8	14.3	1220	2 T06403	resistance complex
3	8	14.3	1240	2 T06404	resistance complex
4	7	12.5	169	2 A64396	hypothetical prote
5	7	12.5	200	2 A70527	hypothetical prote
6	7	12.5	242	2 E83077	conserved hypother
7	7	12.5	250	2 H85067	hypothetical prote
8	7	12.5	319	2 B71312	probable GTP-bindi
9	7	12.5	326	2 H72472	hypothetical prote
10	7	12.5	327	2 E64188	dipeptide transpor
11	7	12.5	390	2 D83057	probable aminotran
12	7	12.5	443	2 T48593	hypothetical prote
13	7	12.5	486	2 F64204	spermidine/putresc
14	7	12.5	491	2 A40283	3',5'-cyclic-nucle
15	7	12.5	519	2 T14783	hypothetical prote
16	7	12.5	530	1 A45334	3',5'-cyclic-nucle
17	7	12.5	586	2 C71695	ctp synthase (pyrG
18	7	12.5	671	2 D38490	maturase-related p
19	7	12.5	673	2 S53604	maturase-related p
20	7	12.5	674	2 T07017	maturase homolog -
21	7	12.5	678	2 B38489	maturase-related p
22	7	12.5	687	2 A34879	maturase-related p
23	7	12.5	688	2 S53605	maturase-related p
24	7	12.5	716	2 G01627	androgen receptor
25	7	12.5	748	2 T30634	hypothetical prote
26	7	12.5	770	1 A44337	kinesin-related pr
27	7	12.5	791	2 T20815	hypothetical prote
28	7	12.5	838	2 S49750	probable membrane
29	7	12.5	974	2 A71466	probable zinc meta

30	7	12.5	1270	2 T30339	dsRNA adenosine de
31	7	12.5	1341	2 S66835	probable membrane
32	6	10.7	42	2 T07291	hypothetical prote
33	6	10.7	59	1 ERBP1A	gene 1A protein -
34	6	10.7	68	1 IABY3	proteinase A inhib
35	6	10.7	79	2 A82875	ribosomal protein
36	6	10.7	81	2 S61847	hrpX protein - Pse
37	6	10.7	87	2 C82334	probable host fact
38	6	10.7	90	2 H89159	hypothetical prote
39	6	10.7	91	2 D64066	host factor I - Ha
40	6	10.7	93	2 T40184	conserved hypother
41	6	10.7	94	2 T12719	hypothetical prote
42	6	10.7	100	2 PH1075	Ig light chain V r
43	6	10.7	109	2 PH0089	Ig kappa chain V r
44	6	10.7	109	2 F82067	probable anti-sigm
45	6	10.7	111	2 S37723	outer surface prot

ALIGNMENTS

RESULT 1

G75638  
nodulation protein-related protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: G75638  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: G75638  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-703 <WHI>  
A:Cross-references: GB:A5001827; NID:g6460959; PIDN:AAF12674.1; PID:g6460971; TIGR:DR  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DRC0037  
A:Map position: plasmid  
A:Genome: plasmid  
A:Note: plasmid CPl

Query Match 14.3%; Score 8; DB 2; Length 703;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AASVLEAT 46  
Db 662 AASVLEAT 669

RESULT 2

T06403  
resistance complex protein I2C-1 - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 17-Nov-2000  
C:Accession: T06403  
R:Ori, N.; Eshed, Y.; Paran, I.; Presting, G.; Aviv, D.; Tanksley, S.; Zamir, D.; Flu  
Plant Cell 9, 521-532, 1997  
A:Title: The I2C family from the wilt disease resistance locus I2 belongs to the nuclei  
A:Reference number: Z15652; MUID:97290204  
A:Accession: T06403  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1220 <ORI>  
A:Cross-references: EMBL:AF004878; NID:g2258314; PIDN:AAB63274.1; PID:g2258315  
C:Genetics:  
A:Gene: I2C-1  
A:Map position: 11

---

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/297,510  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5602019and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-297-510-6

Query Match 12.5%; Score 7; DB 1; Length 530;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVLEA 45  
Db 59 AASVLEA 65

RESULT 15  
US-08-479-532-6  
Sequence 6, Application US/08479532  
Patent No. 5776752  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
TITLE OF INVENTION: Phosphodiesterases  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,532  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/297,494  
FILING DATE:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:

NAME: No. 5776752and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-479-532-6

Query Match 12.5%; Score 7; DB 1; Length 530;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVLEA 45  
Db 59 AASVLEA 65

Search completed: September 15, 2001, 12:47:52  
Job time: 182 sec



```
Qy 39 AASVLEA 45
Db 43 AASVLEA 49

RESULT 12
US-07-872-644-6
; Sequence 6, Application US/07872644
; Patent No. 5389527
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/872.644
; FILING DATE: 19920420
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5389527and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-872-644-6

Query Match 12.5%; Score 7; DB 1; Length 530;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVLEA 45
Db 59 AASVLEA 65

RESULT 13
US-08-297-494-6
; Sequence 6, Application US/08297494
; Patent No. 5580771
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/297,494
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5580771and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-297-494-6

Query Match 12.5%; Score 7; DB 1; Length 530;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVLEA 45
Db 59 AASVLEA 65

RESULT 14
US-08-297-510-6
; Sequence 6, Application US/08297510
; Patent No. 5602019
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/297,494
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5580771and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-297-494-6

Query Match 12.5%; Score 7; DB 1; Length 530;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVLEA 45
Db 59 AASVLEA 65
```

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/297,494  
;; FILING DATE: 04-APR-1991  
;; APPLICATION NUMBER: US 07/688,356  
;; FILING DATE: 04-APR-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. 5800987and, Greta E.  
;; REGISTRATION NUMBER: 35,302  
;; REFERENCE/DOCKET NUMBER: 27866/30822  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (312) 346-5750  
;; TELEFAX: (312) 984-9740  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 17:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 514 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-455-525-17

Query Match 12.5%; Score 7; DB 1; Length 514;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AASVLEA 45  
|||||||  
Db 43 AASVLEA 49

RESULT 10  
US-09-139-491-17  
;; Sequence 17, Application US/09139491  
;; Patent No. 6015677  
;; GENERAL INFORMATION:  
;; APPLICANT: Beavo, Joseph A.  
;; APPLICANT: Bentley, Kelley  
;; APPLICANT: Charbonneau, Harry  
;; APPLICANT: Sonnenburg, William K.  
;; TITLE OF INVENTION: DNA Encoding Mammalian  
;; TITLE OF INVENTION: Phosphodiesterases  
;; NUMBER OF SEQUENCES: 58  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
;; ADDRESSEE: Bicknell  
;; STREET: Two First National Plaza, 20 South Clark  
;; STREET: Street  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60603  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/139,491  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/455,525  
;; FILING DATE: 31-MAY-1995  
;; APPLICATION NUMBER: 08/297,494  
;; FILING DATE:  
;; APPLICATION NUMBER: US 07/688,356  
;; FILING DATE: 04-APR-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. 6015677and, Greta E.  
;; REGISTRATION NUMBER: 35,302  
;; REFERENCE/DOCKET NUMBER: 27866/30822  
;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (312) 346-5750  
;; TELEFAX: (312) 984-9740  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 17:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 514 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-09-139-491-17

Query Match 12.5%; Score 7; DB 3; Length 514;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AASVLEA 45  
|||||||  
Db 43 AASVLEA 49

RESULT 11  
PCT-US92-03222-17  
;; Sequence 17, Application PC/TUS9203222  
;; GENERAL INFORMATION:  
;; APPLICANT: Beavo, Joseph A.  
;; APPLICANT: Bentley, Kelley  
;; APPLICANT: Charbonneau, Harry  
;; APPLICANT: Sonnenburg, William K.  
;; TITLE OF INVENTION: DNA Encoding Mammalian  
;; TITLE OF INVENTION: Phosphodiesterases  
;; NUMBER OF SEQUENCES: 58  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
;; ADDRESSEE: Bicknell  
;; STREET: Two First National Plaza, 20 South Clark  
;; STREET: Street  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60603  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US92/03222  
;; FILING DATE: 19920420  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/688,356  
;; FILING DATE: 04-APR-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Noland, Greta E.  
;; REGISTRATION NUMBER: 35,302  
;; REFERENCE/DOCKET NUMBER: 27866/30822  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (312) 346-5750  
;; TELEFAX: (312) 984-9740  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 17:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 514 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
PCT-US92-03222-17

Query Match 12.5%; Score 7; DB 5; Length 514;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
TITLE OF INVENTION: Phosphodiesterases  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,532  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/297,494  
FILING DATE:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5776752and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-479-532-17

Query Match 12.5% Score 7; DB 1; Length 514;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AASVLEA 45  
| | | | | | |  
Db 43 AASVLEA 49

RESULT 8  
US-08-455-526-17  
Sequence 17, Application US/08455526  
Patent No. 5789553  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
TITLE OF INVENTION: Phosphodiesterases  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago

STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,526  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/297,494  
FILING DATE: 29-AUG-1994  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5789553and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-455-526-17

Query Match 12.5% Score 7; DB 1; Length 514;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 AASVLEA 45  
| | | | | | |  
Db 43 AASVLEA 49

RESULT 9  
US-08-455-525-17  
Sequence 17, Application US/08455525  
Patent No. 5800987  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
TITLE OF INVENTION: Phosphodiesterases  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,525  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-872-644-17

Query Match 12.5%; Score 7; DB 1; Length 514;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVLEA 45  
Db 43 AASVLEA 49

RESULT 5  
US-08-297-494-17  
; Sequence 17, Application US/08297494  
; Patent No. 5580771  
; GENERAL INFORMATION:  
; APPLICANT: Beavo, Joseph A.  
; APPLICANT: Bentley, Kelley  
; APPLICANT: Charbonneau, Harry  
; APPLICANT: Sonnenburg, William K.  
; TITLE OF INVENTION: DNA Encoding Mammalian  
; TITLE OF INVENTION: Phosphodiesterases  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; STREET: Street  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/297,494  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/688,356  
; FILING DATE: 04-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5580771and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/30822  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 514 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-297-494-17

Query Match 12.5%; Score 7; DB 1; Length 514;

Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 39 AASVLEA 45  
Db 43 AASVLEA 49

RESULT 6  
US-08-297-510-17  
; Sequence 17, Application US/08297510  
; Patent No. 5602019  
; GENERAL INFORMATION:  
; APPLICANT: Beavo, Joseph A.  
; APPLICANT: Bentley, Kelley  
; APPLICANT: Charbonneau, Harry  
; APPLICANT: Sonnenburg, William K.  
; TITLE OF INVENTION: DNA Encoding Mammalian  
; TITLE OF INVENTION: Phosphodiesterases  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; STREET: Street  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/297,510  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/688,356  
; FILING DATE: 04-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5602019and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/30822  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 514 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-297-510-17

Query Match 12.5%; Score 7; DB 1; Length 514;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVLEA 45  
Db 43 AASVLEA 49

RESULT 7  
US-08-479-532-17  
; Sequence 17, Application US/08479532  
; Patent No. 5776752  
; GENERAL INFORMATION:  
; APPLICANT: Beavo, Joseph A.

Db 554 EQLRTLTP 561

## RESULT 2

US-08-930-996A-4  
; Sequence 4, Application US/08930996A  
; Patent No. 6100449  
; GENERAL INFORMATION:  
; APPLICANT: FLUHR, Robert  
; APPLICANT: ESHED, Yuval  
; APPLICANT: ORI, Naomi  
; APPLICANT: PARAN, Ilan  
; APPLICANT: ZAMIR, Daniel  
; TITLE OF INVENTION: A GENE FAMILY FROM THE I2 FUSARIUM RESISTANCE  
; TITLE OF INVENTION: LOCUS OF TOMATO AND USE THEREOF FOR TRANSFORMATION AND  
; TITLE OF INVENTION: SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/930,996A

FILING DATE: 09-DEC-1997

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/05272

FILING DATE: 15-APR-1996

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: IL 113,373

FILING DATE: 13-APR-1995

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1240 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-930-996A-4

Query Match 14.3%; Score 8; DB 3; Length 1240;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 EQLRTLTP 30

Db 545 EQLRTLTP 552

## RESULT 3

US-08-630-916A-76  
; Sequence 76, Application US/08630916A  
; Patent No. 601137  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; APPLICANT: Kay, Brian K.  
; APPLICANT: Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States

ZIP: 10036-2711  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/630,916A  
; FILING DATE: 03-APR-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MISROCK, S. LESLIE

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-203

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 896-8864/9741

INFORMATION FOR SEQ ID NO: 76:

SEQUENCE CHARACTERISTICS:

LENGTH: 52 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: Peptide

US-08-630-916A-76

Query Match 12.5%; Score 7; DB 3; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSKEKLR 14

Db 30 SSKEKLR 36

## RESULT 4

US-07-872-644-17  
; Sequence 17, Application US/07872644  
; Patent No. 5389527  
; GENERAL INFORMATION:  
; APPLICANT: Beavo, Joseph A.  
; APPLICANT: Bentley, Kelley  
; APPLICANT: Charbonneau, Harry  
; APPLICANT: Sonnenburg, William K.  
; TITLE OF INVENTION: DNA Encoding Mammalian  
; TITLE OF INVENTION: Phosphodiesterases  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; STREET: Two First National Plaza, 20 South Clark  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/872,644  
; FILING DATE: 19920420  
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/688,356

FILING DATE: 04-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: No. 5389527and, Greta E.

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 27866/30822

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:47:51 : Search time 35.36 Seconds  
(without alignments)  
32.609 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_189  
Perfect score: 56  
Sequence: 1 KISLLHSSKELRRRIKY.....NDAASVLEATVDYVKIREK 56

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgnl\_7/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgnl\_7/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgnl\_7/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgnl\_7/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgnl\_7/ptodata/1/iaa/PCTUS.COMB.pep.\*  
6: /cgnl\_7/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	14.3	1220	3	US-08-930-996A-2
2	8	14.3	1240	3	US-08-930-996A-4
3	7	12.5	52	3	US-08-630-916A-76
4	7	12.5	514	1	US-07-872-644-17
5	7	12.5	514	1	US-08-297-494-17
6	7	12.5	514	1	US-08-297-510-17
7	7	12.5	514	1	US-08-479-532-17
8	7	12.5	514	1	US-08-455-526-17
9	7	12.5	514	3	US-08-455-525-17
10	7	12.5	514	3	US-09-139-491-17
11	7	12.5	514	5	PCT-US92-03222-17
12	7	12.5	530	1	US-07-872-644-6
13	7	12.5	530	1	US-08-297-494-6
14	7	12.5	530	1	US-08-297-510-6
15	7	12.5	530	1	US-08-479-532-6
16	7	12.5	530	1	US-08-455-526-6
17	7	12.5	530	1	US-08-455-525-6
18	7	12.5	530	3	US-09-139-491-6
19	7	12.5	530	5	PCT-US92-03222-6
20	7	12.5	535	1	US-07-872-644-49
21	7	12.5	535	1	US-08-297-494-49
22	7	12.5	535	1	US-08-297-510-49
23	7	12.5	535	1	US-08-479-532-49
24	7	12.5	535	1	US-08-455-526-49
25	7	12.5	535	1	US-08-455-525-49
26	7	12.5	535	3	US-09-139-491-49
27	7	12.5	535	5	PCT-US92-03222-49

28	6	10.7	15	4	US-08-986-837-8	Sequence 8, Appli
29	6	10.7	27	1	US-08-447-702-1	Sequence 1, Appli
30	6	10.7	27	1	US-08-465-615-1	Sequence 1, Appli
31	6	10.7	79	2	US-08-611-510-6	Sequence 6, Appli
32	6	10.7	112	1	US-08-477-877B-87	Sequence 87, Appli
33	6	10.7	112	1	US-08-477-877B-88	Sequence 88, Appli
34	6	10.7	112	2	US-08-472-281A-87	Sequence 87, Appli
35	6	10.7	112	2	US-08-472-281A-88	Sequence 88, Appli
36	6	10.7	112	2	US-08-477-989B-87	Sequence 87, Appli
37	6	10.7	112	2	US-08-477-989B-88	Sequence 88, Appli
38	6	10.7	112	4	US-09-240-274-30	Sequence 30, Appli
39	6	10.7	132	1	US-08-477-877B-84	Sequence 84, Appli
40	6	10.7	132	1	US-08-477-877B-91	Sequence 91, Appli
41	6	10.7	132	2	US-08-472-281A-84	Sequence 84, Appli
42	6	10.7	132	2	US-08-472-281A-91	Sequence 91, Appli
43	6	10.7	132	2	US-08-477-989B-84	Sequence 84, Appli
44	6	10.7	132	2	US-08-477-989B-91	Sequence 91, Appli
45	6	10.7	147	4	US-08-986-837-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-930-996A-2  
; Sequence 2, Application US/08930996A  
; Patent No. 6100449  
; GENERAL INFORMATION:  
; APPLICANT: FLUHR, Robert  
; APPLICANT: ESHED, Yuval  
; APPLICANT: ORI, Naomi  
; APPLICANT: PARAN, Daniel  
; APPLICANT: ZAMIR, Ilan  
; TITLE OF INVENTION: A GENE FAMILY FROM THE I2 FUSARIUM RESISTANCE  
; TITLE OF INVENTION: LOCUS OF TOMATO AND USE THEREOF FOR TRANSFORMATION AND  
; TITLE OF INVENTION: SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/930,996A  
; FILING DATE: 09-DEC-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/05272  
; FILING DATE: 15-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: IL 113,373  
; FILING DATE: 13-APR-1995  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1220 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-930-996A-2

Query Match 14.3% Score 8: DB 3: Length 1220;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 8: Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 23 BQLRTLTP 30  
|||||||

**THIS PAGE BLANK (USPTO)**

---



PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 12.5%; Score 7; DB 21; Length 443;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LRRERIK 19  
 |||||  
 Db 227 lrrerik 233

## RESULT 15

AAAY37305  
 ID AAY37305 standard; Protein; 477 AA.

AC AAY37305;

DT 07-OCT-1999 (first entry)

DE Amino acid sequence of a Chlamydia trachomatis protein.

KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;  
 KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;  
 KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;  
 KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis.

OS Chlamydia trachomatis.

PN WO9928475-A2.

PD 10-JUN-1999.

PF 27-NOV-1998; 98WO-IB01939.

PR 04-NOV-1998; 98US-0107077.

PR 28-NOV-1997; 97FR-0015041.

PR 17-DEC-1997; 97FR-0016034.

XX (GEST ) GENSET.

XX Griffais R;

XX WPI; 1999-371125/31.

XX Genome sequence of Chlamydia trachomatis

XX Disclosure; Page 1042; 1755pp; English.

XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome  
 of Chlamydia trachomatis (see AA201425). The polypeptides can be used as  
 vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences  
 can also be used to control growth of the microorganism. Chlamydia  
 trachomatis is responsible for a large number of diseases, e.g. eye  
 diseases such as conventional trachoma, nonendemic trachoma,  
 paratrachoma, and inclusion conjunctivitis; genital diseases such as  
 nongonococcal urethritis, epididymitis, cervicitis, salpingitis,  
 perihepatitis, Bartholinitis; pneumopathy in breast feeding infants;  
 and venereal lymphogranulomatosis. The polypeptides of the invention  
 may be of use in treating these diseases.

XX Sequence 477 AA;

Query Match 12.5%; Score 7; DB 20; Length 477;

Best Local Similarity 100.0%; Pred. No. 30;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 YVKVIRE 55

|||||

Db 203 yvkviRE 209

Search completed: September 15, 2001, 12:47:10  
 Job time: 160 sec

PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 24-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144894.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.



XX WPI; 2001-071395/08.  
DR  
XX  
PT Secreted and transmembrane proteins and nucleic acids designated PRO,  
PT useful as hybridization probes, in chromosome and gene mapping and gene  
PT therapy -  
XX  
PS Claim 1; Fig 116; 787pp; English.  
XX  
CC The present invention relates to secreted and transmembrane proteins.  
CC These proteins and the DNA encoding them may be used as hybridization  
CC probes, in chromosome and gene mapping and in the generation of  
CC anti-sense RNA and DNA. They may also be used to generate either  
CC transgenic animals or knockout animals which are in turn useful for  
CC development and screening of therapeutically useful reagents.  
CC The nucleic acids may also be used in gene therapy.  
XX  
SQ Sequence 323 AA;  
  
Query Match 12.5%; Score 7; DB 22; Length 323;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 KISLHLS 8  
Db 215 Kislhls 221  
|||||  
  
RESULT 13  
ID AAG15757 standard; Protein; 341 AA.  
XX  
AC AAG15757;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 16134.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
PN  
XX EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127482.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.

PR 24-SEP-1998; 98US-0101915.  
PR 24-SEP-1998; 98US-0101916.  
PR 29-SEP-1998; 98US-0102207.  
PR 29-SEP-1998; 98US-0102240.  
PR 29-SEP-1998; 98US-0102307.  
PR 29-SEP-1998; 98US-0102330.  
PR 29-SEP-1998; 98US-0102331.  
PR 30-SEP-1998; 98US-0102484.  
PR 30-SEP-1998; 98US-0102487.  
PR 30-SEP-1998; 98US-0102570.  
PR 30-SEP-1998; 98US-0102571.  
PR 01-OCT-1998; 98US-0102684.  
PR 01-OCT-1998; 98US-0102687.  
PR 02-OCT-1998; 98US-0102685.  
PR 06-OCT-1998; 98US-0103258.  
PR 06-OCT-1998; 98US-0103258.  
PR 07-OCT-1998; 98US-0103449.  
PR 07-OCT-1998; 98US-0103314.  
PR 07-OCT-1998; 98US-0103315.  
PR 07-OCT-1998; 98US-0103328.  
PR 07-OCT-1998; 98US-0103395.  
PR 07-OCT-1998; 98US-0103396.  
PR 07-OCT-1998; 98US-0103401.  
PR 08-OCT-1998; 98US-0103633.  
PR 08-OCT-1998; 98US-0103678.  
PR 08-OCT-1998; 98US-0103679.  
PR 08-OCT-1998; 98US-0103711.  
PR 14-OCT-1998; 98US-0104257.  
PR 20-OCT-1998; 98US-0104987.  
PR 20-OCT-1998; 98US-0105000.  
PR 20-OCT-1998; 98US-0105002.  
PR 21-OCT-1998; 98US-0105104.  
PR 22-OCT-1998; 98US-0105169.  
PR 22-OCT-1998; 98US-0105266.  
PR 26-OCT-1998; 98US-0105693.  
PR 26-OCT-1998; 98US-0105694.  
PR 27-OCT-1998; 98US-0105807.  
PR 27-OCT-1998; 98US-0105881.  
PR 27-OCT-1998; 98US-0105882.  
PR 27-OCT-1998; 98US-0106062.  
PR 28-OCT-1998; 98US-0106023.  
PR 28-OCT-1998; 98US-0106029.  
PR 28-OCT-1998; 98US-0106030.  
PR 28-OCT-1998; 98US-0106032.  
PR 28-OCT-1998; 98US-0106033.  
PR 28-OCT-1998; 98US-0106178.  
PR 29-OCT-1998; 98US-0106248.  
PR 29-OCT-1998; 98US-0106384.  
PR 29-OCT-1998; 98US-0108500.  
PR 30-OCT-1998; 98US-0106464.  
PR 03-NOV-1998; 98US-0106856.  
PR 03-NOV-1998; 98US-0106902.  
PR 03-NOV-1998; 98US-0106905.  
PR 03-NOV-1998; 98US-0106919.  
PR 03-NOV-1998; 98US-0106932.  
PR 03-NOV-1998; 98US-0106934.  
PR 10-NOV-1998; 98US-0107783.  
PR 17-NOV-1998; 98US-0108775.  
PR 17-NOV-1998; 98US-0108779.  
PR 17-NOV-1998; 98US-0108787.  
PR 17-NOV-1998; 98US-0108788.  
PR 17-NOV-1998; 98US-0108801.  
PR 17-NOV-1998; 98US-0108802.  
PR 17-NOV-1998; 98US-0108806.  
PR 17-NOV-1998; 98US-0108806.  
PR 17-NOV-1998; 98US-0108807.  
PR 17-NOV-1998; 98US-0108867.  
PR 17-NOV-1998; 98US-0108925.  
PR 18-NOV-1998; 98US-0108848.  
PR 18-NOV-1998; 98US-0108849.  
PR 18-NOV-1998; 98US-0108850.  
PR 18-NOV-1998; 98US-0108851.  
PR 18-NOV-1998; 98US-0108852.  
PR 18-NOV-1998; 98US-0108858.  
PR 18-NOV-1998; 98US-0108904.

XX (GETH ) GENENTECH INC.  
PA Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;  
XX WPI; 2000-237871/20.  
XX N-PSDB; AAA37079.  
XX New mammalian DNA sequences encoding transmembrane, receptor or  
PT secreted PRO polypeptides, useful for screening of potential peptide or  
PT small molecule inhibitors of the relevant receptor/ligand interactions  
XX Claim 12; Fig 116; 773pp; English.  
XX AAA37022 to AAA37144 encode the new isolated human transmembrane,  
CC receptor or secreted PRO polypeptides given in AAV99340 to AAV99462. The  
CC transmembrane and receptor PRO proteins can be used for screening of  
CC potential peptide or small molecule inhibitors of the relevant  
CC receptor/ligand interactions. The polypeptides and nucleotide sequences  
CC encoding then have various industrial applications, including uses as  
CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent  
CC PCR primers and hybridisation probes used in the isolation of the PRO  
XX polypeptides from the present invention.  
SQ Sequence 323 AA;

Query Match 12.5%; Score 7; DB 21; Length 323;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KISLHLS 8  
|||  
Db 215 kisllhs 221

RESULT 12  
AAB66146  
ID AAB66146 standard; protein; 323 AA.  
AC AAB66146;  
XX 02-APR-2001 (first entry)  
DT Protein of the invention #58.  
DE Secreted; transmembrane; gene therapy.  
XX Unidentified.  
XX WO200078961-A1.  
PN 28-DEC-2000.  
PD 18-FEB-2000; 2000WO-US04342.  
XX 23-JUN-1999; 99US-0141037.  
PR 20-JUL-1999; 99US-0144758.  
PR 26-JUL-1999; 99US-0145698.  
PR 01-SEP-1999; 99WO-US20111.  
PR 29-OCT-1999; 99US-0162506.  
PR 30-NOV-1999; 99WO-US28313.  
PR 02-DEC-1999; 99WO-US28551.  
PR 16-DEC-1999; 99WO-US30095.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00376.  
XX (GETH ) GENENTECH INC.  
PA Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;  
XX Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;  
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D;  
PI Watanabe CK, Williams PM, Wood WI;

PR	10-AUG-1999;	99US-01481171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151086.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161922.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 12.5% Score 7; DB 21; Length 291;		
Best Local Similarity 100.0%; Pred. No. 19;		
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	13 LRRERIK 19	

Db	75 lrrerik 81	
RESULT 11		
AAAY99397		
ID	AAAY99397 standard; Protein; 323 AA.	
XX	AC	AAAY99397;
XX	AC	AAAY99397;
DT	08-AUG-2000	(first entry)
XX	Human PRO1298 (UNQ666)	amino acid sequence SEQ ID NO:210.
DE	Human:	PRO polypeptide; membrane bound protein; receptor; diagnosis;
KW	transmembrane;	secretion; immunoadhesion; pharmaceutical; screening.
KW	Homo sapiens.	
XX	WO200012708-A2.	
PD	09-MAR-2000.	
XX	01-SEP-1999;	99WO-US20111.
PF	01-SEP-1998;	98US-0098716.
XX	01-SEP-1998;	98US-0098749.
PR	01-SEP-1998;	98US-0098750.
PR	02-SEP-1998;	98US-0098803.
PR	02-SEP-1998;	98US-0098821.
PR	02-SEP-1998;	98US-0098843.
PR	09-SEP-1998;	98US-0099536.
PR	09-SEP-1998;	98US-0099596.
PR	09-SEP-1998;	98US-0099598.
PR	09-SEP-1998;	98US-0099602.
PR	09-SEP-1998;	98US-0099642.
PR	10-SEP-1998;	98US-0099741.
PR	10-SEP-1998;	98US-0099754.
PR	10-SEP-1998;	98US-0099763.
PR	10-SEP-1998;	98US-0099792.
PR	10-SEP-1998;	98US-0099808.
PR	10-SEP-1998;	98US-0099812.
PR	10-SEP-1998;	98US-0099815.
PR	15-SEP-1998;	98US-0099816.
PR	15-SEP-1998;	98US-0100385.
PR	15-SEP-1998;	98US-0100388.
PR	15-SEP-1998;	98US-0100390.
PR	16-SEP-1998;	98US-0100584.
PR	16-SEP-1998;	98US-0100627.
PR	16-SEP-1998;	98US-0100661.
PR	16-SEP-1998;	98US-0100662.
PR	16-SEP-1998;	98US-0100664.
PR	17-SEP-1998;	98US-0100683.
PR	17-SEP-1998;	98US-0100684.
PR	17-SEP-1998;	98US-0100710.
PR	17-SEP-1998;	98US-0100711.
PR	17-SEP-1998;	98US-0100919.
PR	17-SEP-1998;	98US-0100930.
PR	18-SEP-1998;	98US-0100848.
PR	18-SEP-1998;	98US-0100849.
PR	18-SEP-1998;	98US-0101014.
PR	18-SEP-1998;	98US-0101068.
PR	18-SEP-1998;	98US-0101071.
PR	22-SEP-1998;	98US-0101279.
PR	23-SEP-1998;	98US-0101471.
PR	23-SEP-1998;	98US-0101472.
PR	23-SEP-1998;	98US-0101474.
PR	23-SEP-1998;	98US-0101475.
PR	23-SEP-1998;	98US-0101476.
PR	23-SEP-1998;	98US-0101477.
PR	23-SEP-1998;	98US-0101479.
PR	24-SEP-1998;	98US-0101738.
PR	24-SEP-1998;	98US-0101741.
PR	24-SEP-1998;	98US-0101743.

```
Query Match      12.58; Score 7; DB 21; Length 250;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 SSKEKLR 14
        |||||
Db      244 sskekIr 250

RESULT 10
AAG15758
ID  AAG15758 standard; Protein; 291 AA.
AC  AAG15758;
XX
XX
DT  17-OCT-2000 (first entry)
XX
DE  Arabidopsis thaliana protein fragment SEQ ID NO: 16135.
XX
KW  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
KW  termination sequence.
XX
OS  Arabidopsis thaliana.
XX
PN  EP1033405-A2.
XX
XX
PD  06-SEP-2000.
XX
PF  25-FEB-2000; 2000EP-0301439.
XX
PR  25-FEB-1999; 99US-0121825.
PR  05-MAR-1999; 99US-0123180.
PR  09-MAR-1999; 99US-0123548.
PR  23-MAR-1999; 99US-0125788.
PR  25-MAR-1999; 99US-0126284.
PR  29-MAR-1999; 99US-0126785.
PR  01-APR-1999; 99US-0127462.
PR  06-APR-1999; 99US-0128234.
PR  08-APR-1999; 99US-0128714.
PR  16-APR-1999; 99US-0129845.
PR  19-APR-1999; 99US-0130077.
PR  21-APR-1999; 99US-0130449.
PR  23-APR-1999; 99US-0130510.
PR  23-APR-1999; 99US-0130891.
PR  28-APR-1999; 99US-0131449.
PR  30-APR-1999; 99US-0132048.
PR  30-APR-1999; 99US-0132407.
PR  04-MAY-1999; 99US-0132484.
PR  05-MAY-1999; 99US-0132485.
PR  06-MAY-1999; 99US-0132486.
PR  06-MAY-1999; 99US-0132487.
PR  07-MAY-1999; 99US-0132863.
PR  11-MAY-1999; 99US-0134256.
PR  14-MAY-1999; 99US-0134218.
PR  14-MAY-1999; 99US-0134219.
PR  14-MAY-1999; 99US-0134221.
PR  14-MAY-1999; 99US-0134370.
PR  18-MAY-1999; 99US-0134768.
PR  19-MAY-1999; 99US-0134941.
PR  20-MAY-1999; 99US-0135124.
PR  21-MAY-1999; 99US-0135353.
PR  24-MAY-1999; 99US-0135629.
PR  25-MAY-1999; 99US-0136021.
PR  27-MAY-1999; 99US-0136382.
PR  28-MAY-1999; 99US-0136782.
PR  01-JUN-1999; 99US-0137222.
PR  03-JUN-1999; 99US-0137528.
PR  04-JUN-1999; 99US-0137502.
PR  07-JUN-1999; 99US-0137724.
PR  08-JUN-1999; 99US-0138094.
PR  10-JUN-1999; 99US-0138540.

10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139452.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
29-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.
13-JUL-1999; 99US-0143542.
14-JUL-1999; 99US-0143624.
15-JUL-1999; 99US-0144005.
16-JUL-1999; 99US-0144085.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.
21-JUL-1999; 99US-0145088.
22-JUL-1999; 99US-0145085.
22-JUL-1999; 99US-0145087.
22-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-0145145.
23-JUL-1999; 99US-0145218.
23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
02-AUG-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
03-AUG-1999; 99US-0147038.
04-AUG-1999; 99US-0147204.
04-AUG-1999; 99US-0147302.
05-AUG-1999; 99US-0147192.
05-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-0147935.
```

PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140333.  
PR 23-JUN-1999; 99US-0140334.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.



PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 21-OCT-1999; 99US-0160880.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 26-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161982.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162142.

Query Match 12.5% Score 7; DB 21; Length 135;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 SSKEKLR 14  
Db 129 sskekrlr 135  
|||||

RESULT 8  
AAY29197  
ID AAY29197 standard; Protein; 170 AA.  
AC AAY29197;  
XX  
XX  
DT 25-OCT-1999 (first entry)  
DE Amino acid sequence of a virulence factor encoded by ORE26844c.  
KW Human pathogen; virulence polypeptide; virulence factor;  
KW pathogenic infection; Pseudomonas aeruginosa infection.  
XX  
OS Pseudomonas aeruginosa.  
XX WO927129-A1.  
XX  
XX 03-JUN-1999.  
XX  
XX 25-NOV-1998; 98WO-US25247.  
XX  
XX 25-NOV-1997; 97US-0066517.  
XX  
XX (GEHO ) GEN HOSPITAL CORP.  
XX  
XX Ausubel F, Cao H, Drenkard E, Goodman HM, Mahajan-Miklos S;  
PI Rahme LG, Tan M, Tsongalis J;  
XX  
XX WPI; 1999-357851/30.  
XX  
XX Virulence factors useful in developing disease treatments  
PT  
PS Disclosure; Fig 4; 228pp; English.  
XX  
XX The present sequence represents a Pseudomonas aeruginosa polypeptide  
CC sequence. P. aeruginosa is an opportunistic human pathogen present in  
CC soil water and plants. The specification describes virulence polypeptides  
CC and nucleic acid sequence encoding such polypeptides. These sequences  
CC can be used to identify a compound which is capable of decreasing the  
CC expression of a pathogenic virulence factor. Compounds that inhibit  
CC the expression or activity of virulence factor polypeptides can be  
CC used to treat pathogenic infections, especially where the infection  
CC is a P. aeruginosa infection.

CC note: the sequences given in the specification were poorly legible, and  
CC in some instances assumptions were made as to the identity of the  
CC residue; it is therefore possible that the sequence given below is  
CC not entirely correct.  
XX  
SQ Sequence 170 AA;  
Query Match 12.5% Score 7; DB 20; Length 170;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 42 VLEATVD 48  
Db 58 vleatvd 64  
|||||

RESULT 9  
AAG43790  
ID AAG43790 standard; Protein; 250 AA.  
XX  
XX AAG43790;  
XX  
XX 18-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 54774.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 28-APR-1999; 99US-0130891.  
XX 28-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 06-MAY-1999; 99US-0132487.  
XX 07-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.

PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.

```

XX 05-DEC-1997 (first entry)
XX
XX Tomato immunity 2 (fungal resistance) gene product.
XX
XX Tomato; transgenic plant; disease resistance; wilt inducing fungi;
XX amplified fragment length polymorphism; AFLP; Immunity 2; I-2 gene;
XX Fusarium oxysporum; genetic engineering.
XX
XX Lycopersicon esculentum.
XX
XX WO9706259-A2.
XX
XX 20-FEB-1997.
XX
XX 06-AUG-1996; 96WO-EP03480.
XX
XX 07-AUG-1995; 95EP-0401849.
XX
XX (KEYG-) KEYGENE NV.
XX
XX Simons G, Vos P, Zabeau M;
XX
XX WPI; 1997-154265/14.
XX
XX N-PSDB; AAT79882.
XX
XX New immunity-2 resistance gene, imparting pathogen resistance to
XX plants - used to transform plants, esp. to protect against
XX wilt-inducing fungi
XX
XX Claim 41; Fig 6a-d; 6lpp; English.
XX
XX AAM25157 shows the product of the tomato immunity 2 (I-2) gene. This
XX gene was used to produce transgenic plants that are resistant to
XX wilt-inducing fungi, e.g. Fusarium 2, especially F. oxysporum f.sp.
XX lycopersici race 2. Plants that may be transformed include tomato,
XX melon, tobacco, Arabidopsis, aubergine and potato.
XX
XX Sequence 1266 AA;

Query Match 14.3%; Score 8; DB 18; Length 1266;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 EQLRTLTP 30
Db 547 eqlrtltp 554
|||||||

RESULT 6
AAB07754
ID AAB07754 standard; Protein; 1266 AA.
AC AAB07754;
XX
XX 07-NOV-2000 (first entry)
XX
XX Amino acid sequence of the protein encoded by the I-2 resistance gene.
XX
XX Regulatory activity; transcription; I-2 resistance gene; tomato;
XX egg plant; potato; melon; tobacco; Arabidopsis; plant pathogen; fungi;
XX tissue-specific.
XX
XX Fusarium oxysporum.
XX
XX EP1024196-A1.
XX
XX 02-AUG-2000.
XX
XX 29-JAN-1999; 99EP-0400212.
XX
XX 29-JAN-1999; 99EP-0400212.

XX (KEYG-) KEYGENE NV.
XX
XX Haring MA, Cornelissen BJC, Mes JJ, Simons AFM;
XX
XX WPI; 2000-516034/47.
XX
XX N-PSDB; AAA59332.
XX
XX New I-2 resistance gene tissue-specific regulatory sequence useful in
XX plant resistance mechanisms against plant pathogens such as fungi -
XX
XX Disclosure; Page 26-31; 47pp; English.
XX
XX The present sequence represents I-2 resistance protein. The specification
XX describes nucleotide sequences which have a regulatory activity on the
XX transcription of the I-2 resistance gene in plant host cells.
XX
XX The transgenic plants, especially tomato, egg plant, potato, melon,
XX tobacco and Arabidopsis, are capable of expressing a gene mediating
XX resistance to a plant pathogen, such as fungi, in a tissue-specific
XX manner. The plant is capable of preventing infection by a plant
XX pathogen, such as fungi. Inserting the regulatory activity polynucleotide
XX into plant cell genomes is useful for providing plants with reduced
XX susceptibility to plant pathogens, especially for protecting plants
XX in cultivation.
XX
XX Sequence 1266 AA;

Query Match 14.3%; Score 8; DB 21; Length 1266;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 EQLRTLTP 30
Db 547 eqlrtltp 554
|||||||

RESULT 7
AAG43791
ID AAG43791 standard; Protein; 135 AA.
AC AAG43791;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 54775.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX
XX 05-MAR-1999; 99US-0123180.
XX
XX 09-MAR-1999; 99US-0123548.
XX
XX 23-MAR-1999; 99US-0125788.
XX
XX 25-MAR-1999; 99US-0126264.
XX
XX 29-MAR-1999; 99US-0126785.
XX
XX 01-APR-1999; 99US-0127452.
XX
XX 06-APR-1999; 99US-0128234.
XX
XX 08-APR-1999; 99US-0128714.
XX
XX 16-APR-1999; 99US-0129845.
XX
XX 19-APR-1999; 99US-0130077.
XX
XX 21-APR-1999; 99US-0130449.
XX
XX 23-APR-1999; 99US-0130510.
XX
XX 23-APR-1999; 99US-0130891.

```

```

FT Domain 646..1220
XX /note= "Leucine-rich repeat region"
PN WO9632007-A1.
XX
XX 17-OCT-1996.
XX
XX 15-APR-1996; 96WO-US05272.
XX
XX 13-APR-1995; 95IL-0113373.
XX
XX (YEDA ) YEDA RES & DEV CO LTD.
XX (YISS ) YISSUM RES & DEV CO.
XX
XX Eshed Y, Fluhr R, Ori N, Paran I, Zamir D;
XX
XX WPI; 1996-476739/47.
XX N-PSDB; AAT42134.
XX
XX New DNA conferring resistance to Fusarium sp. - used for producing
XX disease-resistant tomato plants and for identifying new resistance
XX genes and diseases
XX
XX Claim 1; Fig 4A-C; 50pp; English.
XX
XX This I2C-1 protein is encoded by a sequence from the I2C multigene
XX family from the I2 Fusarium wilt resistance locus of tomato, and
XX confers resistance to Fusarium oxysporum f.sp. lycopersici race-2.
XX This sequence and I2C-2 (AAW03665) are encoded by genes from a locus
XX completely linked to I2, and show structural similarity with other
XX resistance proteins. The protein has a conserved N-terminal
XX nucleotide-binding domain (the P-loop) and 5 other conserved
XX domains of unknown function. At least half the C-terminus is
XX composed of leucine-rich repeats, which may be responsible for
XX specificity of interaction, either with a pathogen protein
XX component, or with downstream factors involved with signal
XX transduction. There does not appear to be a transmembrane domain,
XX indicating an intracellular location. A putative leucine zipper
XX domain has been predicted. I2C genes may be inserted in a cosmid
XX vector for expression in a tomato transgenic plant, to confer
XX disease-resistance, or may be used as restriction fragment length
XX polymorphism probes for screening for and selective breeding of
XX tomato or Solanaceae plants with disease-resistance.
XX
XX Sequence 1220 AA;

Query Match 14.3%; Score 8; DB 17; Length 1220;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 23 EQLRTLTP 30
Db 554 eqlrtltp 561
|||||||

RESULT 4
AAW03665
ID AAW03665 standard; Protein; 1240 AA.
XX
XX AAW03665;
XX
XX 22-FEB-1997 (first entry)
XX
XX I2C-2 protein conferring Fusarium wilt disease-resistance.
XX
XX Tomato; I2C-2; Fusarium wilt; disease-resistance; Fusarium oxysporum;
XX P-loop; leucine-rich repeat; transgenic plant; screening;
XX restriction fragment length polymorphism; crop improvement;
XX Solanaceae.
XX
XX Lycopersicon esculentum.
XX

```

```

PH Key Location/Qualifiers
FT Domain 201..208
XX /note= "P-loop"
FT Domain 273..282
XX /note= "Conserved motif of unknown function"
FT Domain 304..311
XX /note= "Conserved motif of unknown function"
FT Domain 367..372
XX /note= "Conserved motif of unknown function"
FT Misc-difference 387..388
XX /note= "Conserved motif of unknown function"
FT Domain 410..415
XX /note= "Conserved motif of unknown function"
FT Domain 488..497
XX /note= "Conserved motif of unknown function"
XX
XX WO9632007-A1.
XX
XX 17-OCT-1996.
XX
XX 15-APR-1996; 96WO-US05272.
XX
XX 13-APR-1995; 95IL-0113373.
XX
XX (YEDA ) YEDA RES & DEV CO LTD.
XX (YISS ) YISSUM RES & DEV CO.
XX
XX Eshed Y, Fluhr R, Ori N, Paran I, Zamir D;
XX
XX WPI; 1996-476739/47.
XX N-PSDB; AAT42135.
XX
XX New DNA conferring resistance to Fusarium sp. - used for producing
XX disease-resistant tomato plants and for identifying new resistance
XX genes and diseases
XX
XX Claim 1; Fig 4A-C; 50pp; English.
XX
XX This I2C-2 protein is encoded by a sequence from the I2C multigene
XX family from the I2 Fusarium wilt resistance locus of tomato, and
XX confers resistance to Fusarium oxysporum f.sp. lycopersici race-2.
XX This sequence and I2C-1 (AAW03664) are encoded by genes from a locus
XX completely linked to I2, and show structural similarity with other
XX resistance proteins. The protein has a conserved N-terminal
XX nucleotide-binding domain (the P-loop) and 5 other conserved
XX domains of unknown function. At least half the C-terminus is
XX composed of leucine-rich repeats, which may be responsible for
XX specificity of interaction, either with a pathogen protein
XX component, or with downstream factors involved with signal
XX transduction. There does not appear to be a transmembrane domain,
XX indicating an intracellular location. I2C genes may be inserted
XX in a cosmid vector for expression in a tomato transgenic plant, to
XX confer disease-resistance, or may be used as restriction fragment
XX length polymorphism probes for screening for and selective breeding
XX of tomato or Solanaceae plants with disease-resistance.
XX
XX Sequence 1240 AA;

Query Match 14.3%; Score 8; DB 17; Length 1240;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 23 EQLRTLTP 30
Db 545 eqlrtltp 552
|||||||

RESULT 5
AAW25157
ID AAW25157 standard; Protein; 1266 AA.
XX
XX AAW25157;
XX

```

PA (AFAR/) AFAR D E.  
 PA (HUBE/) HUBERT R S.  
 PA (RAIT/) RAITANO A B.  
 XX  
 PI Afar DE, Hubert RS, Raitano AB;  
 XX  
 DR WPI: 2000-237872/20.  
 DR N-PSDB: AAZ94275.  
 XX  
 PT Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -  
 XX  
 PS Claim 1; Fig 2A-D; 62pp; English.  
 XX  
 CC This sequence is that of human PHELIx, a novel basic Helix Loop  
 CC Helix protein thought to act as a transcription factor. PHELIx  
 CC normally exhibits a testis-specific expression pattern but is  
 CC up-regulated in prostate and other types of cancer. The invention  
 CC provides diagnostic and therapeutic methods useful in the  
 CC management of various cancers which express PHELIx, including  
 CC prostate cancer, bladder cancer, ovarian cancer and testicular  
 CC cancer, including therapies aimed at inhibition the transcription,  
 CC translation, processing or function of PHELIx. The expression  
 CC pattern of PHELIx suggests that is an ideal target for a cancer  
 CC vaccine approach to prostate cancer. PHELIx protein can also be  
 CC used to screen for agonists and antagonists of therapeutic value  
 CC and to raise antibodies.  
 XX  
 SQ Sequence 405 AA;

Query Match 100.0%; Score 56; DB 21; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-49;  
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKISLHSSKEKLRRERIKYCCQLRTLLPYVKGKNDAAVLEATVDYVKYIREK 56  
 Db 134 KKISLHSSKEKLRRERIKYCCQLRTLLPYVKGKNDAAVLEATVDYVKYIREK 189  
 |||||  
 |||||

RESULT 2  
 AAY79270  
 ID AAY79270 standard; Peptide; 15 AA.  
 XX  
 AC AAY79270;  
 XX  
 DT 03-JUL-2000 (first entry)  
 XX  
 DE PHELIx peptide used to raise antibody.  
 XX  
 KW PHELIx; human; testis-specific; transcription factor;  
 KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
 KW therapy; diagnosis; vaccine; antibody.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200012709-A2.  
 XX  
 PD 09-MAR-2000.  
 XX  
 PF 31-AUG-1999; 99WO-US20137.  
 XX  
 PR 31-AUG-1998; 98US-0098610.  
 PR 31-OCT-1998; 98US-0106524.  
 XX  
 PA (UROC-) UROGENESIS INC.  
 PA (AFAR/) AFAR D E.  
 PA (HUBE/) HUBERT R S.  
 PA (RAIT/) RAITANO A B.  
 XX  
 PI Afar DE, Hubert RS, Raitano AB;  
 XX

DR WPI: 2000-237872/20.  
 XX  
 PT Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -  
 XX  
 PS Example 5; Page 36; 62pp; English.  
 XX  
 CC The present sequence is that of a peptide derived from human  
 CC PHELIx (see AAY79269), a novel transcription factor that is  
 CC normally expressed only in testis tissue, but which is up-regulated  
 CC in prostate and some other cancers. The peptide was conjugated to  
 CC keyhole limpet haemocyanin and used to raise polyclonal antiserum  
 CC in rabbit. The antiserum demonstrated specificity for PHELIx and  
 CC may therefore be useful for assessing the expression of PHELIx in  
 CC patient samples.  
 XX  
 SQ Sequence 15 AA;

Query Match 25.0%; Score 14; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 HSSKEKLRRERIKY 20  
 Db 1 hsskeklrreriky 14  
 |||||  
 |||||

RESULT 3  
 AAW03664  
 ID AAW03664 standard; Protein; 1220 AA.  
 XX  
 AC AAW03664;  
 XX  
 DT 22-FEB-1997 (first entry)  
 XX  
 DE I2C-1 protein conferring Fusarium wilt disease-resistance.  
 XX  
 KW Tomato; I2C-1; Fusarium wilt; disease-resistance; Fusarium oxysporum;  
 KW P-loop; leucine zipper; leucine-rich repeat; transgenic plant;  
 KW restriction fragment length polymorphism; screening;  
 KW crop improvement; Solanaceae.  
 XX  
 OS Lycopersicon esculentum.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 66..71  
 FT /note= "Repeat sequence"  
 FT Region 103..108  
 FT /note= "Repeat sequence"  
 FT Domain 200..207  
 FT /note= "P-loop"  
 FT Region 256..267  
 FT /note= "Repeat sequence"  
 FT Region 269..280  
 FT /note= "Repeat sequence"  
 FT Domain 286..295  
 FT /note= "Conserved motif of unknown function"  
 FT Domain 317..324  
 FT /note= "Conserved motif of unknown function"  
 FT Domain 380..385  
 FT /note= "Conserved motif of unknown function"  
 FT Misc-difference 400..401  
 FT /note= "Conserved motif of unknown function"  
 FT Domain 419..430  
 FT /note= "Conserved motif of unknown function"  
 FT Domain 498..506  
 FT /note= "Conserved motif of unknown function"  
 FT Domain 559..623  
 FT /note= "Leucine-rich repeat region"  
 FT Domain 624..645  
 FT /note= "Putative leucine zipper domain"

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:47:08 ; Search time 65.58 Seconds  
(without alignments)  
51.768 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_189

Perfect score: 56

Sequence: 1 KISLLHSSKELRRIRIKY.....NDAASVLEATVDVVKYIREK 56

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_0601.\*

- 1: /cgnl\_9/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /cgnl\_9/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /cgnl\_9/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /cgnl\_9/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /cgnl\_9/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /cgnl\_9/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /cgnl\_9/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /cgnl\_9/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /cgnl\_9/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /cgnl\_9/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /cgnl\_9/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /cgnl\_9/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /cgnl\_9/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /cgnl\_9/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /cgnl\_9/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /cgnl\_9/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /cgnl\_9/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /cgnl\_9/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /cgnl\_9/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /cgnl\_9/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /cgnl\_9/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	100.0	405	21	AAV79269 Human testis-speci
2	14	25.0	15	21	AAV79270 PHELIIX peptide use
3	8	14.3	1220	17	AAW03664 I2C-1 protein conf
4	8	14.3	1240	17	AAW03665 I2C-2 protein conf
5	8	14.3	1266	18	AAW25157 Tomato immunity 2
6	8	14.3	1266	21	AAW07754 Amino acid sequenc
7	7	12.5	135	21	AAW43791 Arabidopsis thalia
8	7	12.5	170	20	AAV29197 Amino acid sequenc
9	7	12.5	250	21	AAW43790 Arabidopsis thalia
10	7	12.5	291	21	AAW15758 Arabidopsis thalia
11	7	12.5	323	21	AAV99397 Human PRO1298 (UNQ

12	7	12.5	323	22	AAW66146	Protein of the inv
13	7	12.5	341	21	AAW15757	Arabidopsis thalia
14	7	12.5	443	21	AAW15756	Arabidopsis thalia
15	7	12.5	477	20	AAV37305	Amino acid sequenc
16	7	12.5	514	13	AAW28401	Bovine lung 59 kD
17	7	12.5	514	16	AAW69715	Cyclic-GMP stimula
18	7	12.5	514	18	AAW18037	Bovine lung 59 kDa
19	7	12.5	514	18	AAW11240	Calcium/calmodulin
20	7	12.5	514	19	AAW77037	Bovine lung Ca2+/c
21	7	12.5	514	19	AAW71221	59 kDa CaM-PDE fro
22	7	12.5	514	19	AAW60749	59 kDa bovine CaM-
23	7	12.5	514	21	AAW80972	Bovine lung 59 kD
24	7	12.5	530	13	AAW28395	Bovine brain Cam P
25	7	12.5	530	16	AAW69711	Cyclic-GMP stimula
26	7	12.5	530	18	AAW18036	Bovine brain 61 kD
27	7	12.5	530	18	AAW11243	61 kD brain calciu
28	7	12.5	530	19	AAW77036	Bovine brain Ca2+/
29	7	12.5	530	19	AAW71220	61 kDa CaM-PDE DNA
30	7	12.5	530	19	AAW60747	61 kDa bovine CaM-
31	7	12.5	530	21	AAW80968	Bovine brain 61 kD
32	7	12.5	535	13	AAW28411	Protein encoded by
33	7	12.5	535	16	AAW69731	Cyclic-GMP stimula
34	7	12.5	535	18	AAW18039	Human brain 61 kDa
35	7	12.5	535	18	AAW11255	Hippocampus calciu
36	7	12.5	535	19	AAW71227	Human 61 kDa CaM-p
37	7	12.5	535	19	AAW77043	Human Ca2+/calmodu
38	7	12.5	535	19	AAW60755	Amino acid sequenc
39	7	12.5	535	21	AAW80988	Human 61 kD CaM-PD
40	6	10.7	27	15	AAW15385	Comamonas testoste
41	6	10.7	91	21	AAW03544	Human secreted pro
42	6	10.7	112	20	AAV30183	Amino acid sequenc
43	6	10.7	112	20	AAV30184	Amino acid sequenc
44	6	10.7	125	20	AAV12386	Human 5' EST seque
45	6	10.7	125	21	AAW14135	Bordetella pertuss

#### ALIGNMENTS

RESULT	1
AAV79269	AAV79269 standard; Protein; 405 AA.
ID	AAV79269 standard; Protein; 405 AA.
XX	AC AAV79269;
XX	DT 03-JUL-2000 (first entry)
XX	Human testis-specific transcription factor PHELIIX.
DE	PHELIIX; human; testis-specific; transcription factor;
KW	prostate cancer; bladder cancer; ovary cancer; testicular cancer;
KW	therapy; diagnosis; vaccine.
OS	Homo sapiens.
XX	Key Location/Qualifiers
FT	Peptide 134..150
FT	/note= "nuclear localization signal"
FT	Peptide 163..169
FT	/note= "nuclear localization signal"
FT	Domain 140..189
FT	/note= "basic Helix-Loop-Helix domain"
XX	WO200012709-A2.
PN	09-MAR-2000.
XX	31-AUG-1999; 99WO-US20137.
XX	31-AUG-1998; 98US-0098610.
PR	31-OCT-1998; 98US-0106524.
XX	(UROC-) UROGENESYS INC.
PA	



RA Shinnick T.M.;  
 RT "The 65-kilodalton antigen of Mycobacterium tuberculosis.";  
 RL J. Bacteriol. 169:1080-1088(1987).  
 DR EMBL; M15467; AAA88239.1; -.  
 SQ SEQUENCE 42 AA; 4437 MW; 6A1521B60DBC0E29 CRC64;

Query Match 8.9%; Score 5; DB 2; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVL 43  
 Db 11 AASVL 15

## RESULT 14

O77577 PRELIMINARY; PRT; 42 AA.  
 AC O77577;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE UROCORTIN PRECURSOR (FRAGMENT).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cepoi D., Sutton S., Vale W.W.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 14-42 FROM N.A.  
 RA Baigent S.M., Lowry P.J.;  
 RT "The cloning of ovine urocortin.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF051807; AAC27288.1; -.  
 DR EMBL; AF084258; AAC33478.1; -.  
 DR InterPro; IPR000187; -.  
 DR Pfam; PF00473; CRF; 1.  
 DR PROSITE; PS00511; CRF; 1.  
 DR SMART; SM00039; CRF; 1.  
 ET NON\_TER 1  
 SQ SEQUENCE 42 AA; 4893 MW; 31D6139D4937C9E2 CRC64;

Query Match 8.9%; Score 5; DB 6; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29  
 Db 14 LRTLL 18

## RESULT 15

O20145 PRELIMINARY; PRT; 42 AA.  
 AC O20145;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE ORF42A.  
 OS Chlorella vulgaris.  
 OC Chloroplast.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
 OC Chlorellaceae; Chlorella.  
 OX NCBI\_TaxID=3077;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97303241; PubMed=9159184;

RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,  
 RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,  
 RA Inamura A., Yoshinaga K., Sugiura M.;  
 RT "Complete nucleotide sequence of the chloroplast genome from the green  
 RT alga Chlorella vulgaris: the existence of genes possibly involved in  
 RT chloroplast division.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
 DR EMBL; AB001684; BAA57895.1; -.  
 KW Chloroplast.  
 SQ SEQUENCE 42 AA; 5252 MW; 445B221534CACD71 CRC64;

Query Match 8.9%; Score 5; DB 8; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 SKEKL 13  
 Db 19 SKEKL 23

Search completed: September 15, 2001, 12:54:08  
 Job time: 107 sec



"Sequence analysis of the genome of the unicellular cyanobacterium  
 RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 DR EMBL; D64003; BAA10494.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 39 AA; 4659 MW; EBD9BE03539EAF2B CRC64;

Query Match 8.9%; Score 5; DB 2; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KISLL 6  
 Db 9 KISLL 13

RESULT 10  
 Q9RCR8 PRELIMINARY; PRT; 39 AA.  
 ID Q9RCR8  
 AC Q9RCR8  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE M-PROTEIN (FRAGMENT).  
 GN M.  
 OS Streptococcus equi.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 OC NCBI\_TaxID=1336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3;  
 RA Chanter N., Talbot N., Newton R., Verheven K.;  
 RT "Independent multiple emergence of Streptococcus equi with truncated  
 RT M-like proteins in long-term carriers from different outbreaks of  
 RT equine strangles.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ249870; CAB64607.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 39  
 SQ SEQUENCE 39 AA; 4016 MW; 0C7AC9D7AC21AF60 CRC64;

Query Match 8.9%; Score 5; DB 2; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVL 43  
 Db 2 AASVL 6

RESULT 11  
 O19688 PRELIMINARY; PRT; 39 AA.  
 ID O19688  
 AC O19688  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE HLA-B\*27 VARIANT EXON 2 (ALPHA DOMAIN) (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Blaszyk R., Weber M., Salama A.;  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X83727; CAA58698.1; -.  
 DR HSSP; P01891; 2HLA.  
 DR InterPro; IPR001039; -.

DR Pfam; PF00129; MHC\_I; 1.  
 KW MHC.  
 FT NON\_TER 1  
 FT NON\_TER 39  
 SQ SEQUENCE 39 AA; 4748 MW; F5E3FD2A53138908 CRC64;

Query Match 8.9%; Score 5; DB 7; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29  
 Db 27 LRTLL 31

RESULT 12  
 Q9KKK8 PRELIMINARY; PRT; 40 AA.  
 ID Q9KKK8  
 AC Q9KKK8  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE HYPOTHETICAL PROTEIN VCA1064.  
 GN VCA1064.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OC NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae.";  
 RL Nature 406:477-483(2000).  
 DR EMBL; AE004432; AAF96958.1; -.  
 DR TIGR; VCA1064; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 40 AA; 4465 MW; 9861C2D6CBF96BAF CRC64;

Query Match 8.9%; Score 5; DB 2; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 KGRKN 37  
 Db 35 KGRKN 39

RESULT 13  
 Q50817 PRELIMINARY; PRT; 42 AA.  
 ID Q50817  
 AC Q50817  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE 65 KDA ANTIGEN (CELL WALL PROTEIN A) GENE (CELL WALL PROTEIN A).  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OC NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ERDMAN;  
 RX MEDLINE=87137260; PubMed=3029018;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29  
| | | | |  
Db 16 LRTLL 20

RESULT 6  
Q9W7P0 PRELIMINARY; PRT; 31 AA.  
AC Q9W7P0;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE CLASS 7A MYOSIN (FRAGMENT).  
OS Morone saxatilis (Striped bass).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidel;  
OC Moronidae; Morone.  
OX NCBI\_TaxID=34816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=RPE;  
RA Hillman D.W., Bost-Usinger L., Cheng J., Burnside B.;  
RT "Multiple Myosins are Expressed in Fish RPE and Retina.";  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF002867; AA041448.1; -.  
DR InterPro; IPR001609; -.  
DR Pfam; PF00063; myosin\_head; 1.  
DR ProDom; PD000355; -. 1.  
FT NON\_TER 1  
FT NON\_TER 31  
SQ SEQUENCE 31 AA; 3465 MW; 594C38532AB69B04 CRC64;

Query Match 8.9%; Score 5; DB 13; Length 31;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 VLEAT 46  
| | | | |  
Db 24 VLEAT 28

RESULT 7  
Q9RCT4 PRELIMINARY; PRT; 35 AA.  
AC Q9RCT4;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE M-PROTEIN (FRAGMENT).  
GN M.  
OS Streptococcus equi.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1336;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=E;  
RA Chanter N., Talbot N., Newton R., Verheyen K.;  
RT "Independent multiple emergence of Streptococcus equi with truncated  
RT M-like proteins in long-term carriers from different outbreaks of  
RT equine strangles.";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ249874; CAB64611.1; -.  
FT NON\_TER 1  
FT NON\_TER 35  
SQ SEQUENCE 35 AA; 3540 MW; 925BAAA2F0C45DC4 CRC64;

Query Match 8.9%; Score 5; DB 2; Length 35;

Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVL 43  
| | | | |  
Db 2 AASVL 6

RESULT 8  
Q9R5H9 PRELIMINARY; PRT; 35 AA.  
AC Q9R5H9;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE SURFACE ARRAY PROTEIN (FRAGMENT).  
OS Aeromonas hydrophila.  
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;  
OC Aeromonas.  
OX NCBI\_TaxID=644;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92407495; PubMed=1382113;  
RA Kokka R.P., Vedros N.A., Janda J.M.;  
RT "Immunochemical analysis and possible biological role of an Aeromonas  
RT hydrophila surface array protein in septicemia.";  
RL J. Gen. Microbiol. 138:1229-1236(1992).  
SQ SEQUENCE 35 AA; 3537 MW; 2E8BDDC978B4C CRC64;

Query Match 8.9%; Score 5; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 DAASV 42  
| | | | |  
Db 5 DAASV 9

RESULT 9  
P74773 PRELIMINARY; PRT; 39 AA.  
AC P74773;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL 4.7 KDA PROTEIN.  
GN SGL0001.  
OS Synecocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PCC6803;  
RA Tabata S.;  
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96127529; PubMed=8590279;  
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
RA Sugiyama M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
RL region from map positions 64% to 92% of the genome.";  
RNA Res. 2:153-166(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hiroseawa M., Sugiyama M., Nakazaki N., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Okumura S.,  
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;

Db 17 SSKXL 22

## RESULT 2

O20166  
 ID O20166 PRELIMINARY; PRT; 42 AA.  
 AC O20166;  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE ORF42C.  
 OS Chlorella vulgaris.  
 OG Chloroplast.  
 OC Eukaryota: Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
 OC Chlorellaceae; Chlorella.  
 OX NCBI\_TaxID=3077;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97303241; PubMed=9159184;  
 RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,  
 RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,  
 RA Inamura A., Yoshinaga K., Sugiura M.;  
 RT "Complete nucleotide sequence of the chloroplast genome from the green  
 alga *Chlorella vulgaris*: the existence of genes possibly involved in  
 chloroplast division";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
 DR EMBL; AB001684; BAA57939.1; -.  
 KW Chloroplast.  
 SQ SEQUENCE 42 AA; 5355 MW; 136FF356DC528715 CRC64;

Query Match 10.7%; Score 6; DB 8; Length 42;

Best Local Similarity 100.0%; Pred. No. 54;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKISLL 6  
 |||||  
 Db 34 KKISLL 39

## RESULT 3

O9RGT6  
 ID O9RGT6 PRELIMINARY; PRT; 29 AA.  
 AC O9RGT6;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE M-PROTEIN (FRAGMENT).  
 GN M.  
 OS Streptococcus equi.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OX NCBI\_TaxID=1336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=C;  
 RA Chanter N., Talbot N., Newton R., Verheyen K.;  
 RT "Independent multiple emergence of *Streptococcus equi* with truncated  
 M-like proteins in long-term carriers from different outbreaks of  
 equine strangles";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ249872; CAB64609.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 29  
 SQ SEQUENCE 29 AA; 2860 MW; EEFBE55184CCIFE7 CRC64;

Query Match 8.9%; Score 5; DB 2; Length 29;

Best Local Similarity 100.0%; Pred. No. 4.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AASVL 43  
 |||||

Db 2 AASVL 6

## RESULT 4

O9RGT5  
 ID O9RGT5 PRELIMINARY; PRT; 31 AA.  
 AC O9RGT5;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE M-PROTEIN (FRAGMENT).  
 GN M.  
 OS Streptococcus equi.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OX NCBI\_TaxID=1336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=D;  
 RA Chanter N., Talbot N., Newton R., Verheyen K.;  
 RT "Independent multiple emergence of *Streptococcus equi* with truncated  
 M-like proteins in long-term carriers from different outbreaks of  
 equine strangles";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ249873; CAB64610.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 31  
 SQ SEQUENCE 31 AA; 3061 MW; C06E2E6898306D01 CRC64;

Query Match 8.9%; Score 5; DB 2; Length 31;

Best Local Similarity 100.0%; Pred. No. 4.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 39 AASVL 43  
 |||||  
 Db 2 AASVL 6

## RESULT 5

O9KDP6  
 ID O9KDP6 PRELIMINARY; PRT; 31 AA.  
 AC O9KDP6;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
 DE HYPOTHETICAL PROTEIN VC2477.  
 GN VC2477.  
 OS *Vibrio cholerae*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=EL TOR N16961 / SEROTYPE O1;  
 MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*  
*cholerae*";  
 RL Nature 406:477-483(2000).  
 DR EMBL; AE004317; AAF95619.1; -.  
 DR TIGR; VC2477; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 31 AA; 3493 MW; 33DC6391D3FBDC4F CRC64;

Query Match 8.9%; Score 5; DB 2; Length 31;

Best Local Similarity 100.0%; Pred. No. 4.5e+02;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:52:21 ; Search time 17.51 Seconds  
(without alignments)  
423.134 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_189

Perfect score: 56

Sequence: 1 KXISLHSSKEKLRRRIKY .....NDAASVLEATVDYVKVIREK 56

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 30047

Minimum DB seq length: 0

Maximum DB seq length: 56

Post-processing: Listing first 45 summaries

Database :

SPTREMBL16:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_unclassified:\*
- 13: sp\_vertebrate:\*
- 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	10.7	39	Q9UBN9	Q9ubn9 homo sapien
2	6	10.7	42	O20166	O20166 chlorella v
3	5	8.9	29	Q9RCT6	Q9rct6 streptococc
4	5	8.9	31	Q9RCT5	Q9rct5 streptococc
5	5	8.9	31	Q9KP96	Q9kpn6 vibrio chol
6	5	8.9	31	Q9W7P0	Q9w7p0 morone saxa
7	5	8.9	35	Q9RCT4	Q9rct4 streptococc
8	5	8.9	35	Q9RSH9	Q9rsh9 aeromonas h
9	5	8.9	39	P74773	P74773 synechocyst
10	5	8.9	39	Q9RCT8	Q9rct8 streptococc
11	5	8.9	39	Q9KKN8	Q9kkn8 homo sapien
12	5	8.9	40	Q9KKN8	Q9kkn8 vibrio chol
13	5	8.9	42	Q50817	Q50817 mycobacteri
14	5	8.9	42	Q77577	Q77577 ovis aries
15	5	8.9	42	O20145	O20145 chlorella v
16	5	8.9	43	Q9KLS7	Q9kls7 chlamydia p
17	5	8.9	46	Q9FSV1	Q9fsv1 fagus sylv
18	5	8.9	47	Q9UJ31	Q9uj31 schizosacch
19	5	8.9	47	O20181	O20181 chlorella v

20	5	8.9	47	14	Q9QCM2	Q9qcm2 borna disea
21	5	8.9	47	14	Q9QCL9	Q9qcl9 borna disea
22	5	8.9	47	14	Q9QCL6	Q9qcl6 borna disea
23	5	8.9	47	14	Q9QCL3	Q9qcl3 borna disea
24	5	8.9	47	14	Q9QCL0	Q9qcl0 borna disea
25	5	8.9	47	14	Q9QCK7	Q9qck7 borna disea
26	5	8.9	47	14	Q9QCK4	Q9qck4 borna disea
27	5	8.9	47	14	Q9QCK1	Q9qck1 borna disea
28	5	8.9	50	2	Q9FBJ5	Q9fbj5 streptomyce
29	5	8.9	53	1	O28788	O28788 archaeoglob
30	5	8.9	53	2	Q9PGC6	Q9pgc6 xyliella fas
31	5	8.9	53	14	P88748	P88748 human immun
32	5	8.9	54	2	Q9P9S9	Q9p9s9 xyliella fas
33	5	8.9	55	2	Q9RCT7	Q9rct7 streptococc
34	5	8.9	55	4	Q9UGF2	Q9ugf2 homo sapien
35	5	8.9	55	4	Q9NTY5	Q9nty5 homo sapien
36	5	8.9	56	14	Q9Q5B5	Q9q5b5 human immun
37	4	7.1	9	11	Q9QWC2	Q9qwg2 mus musculu
38	4	7.1	11	3	Q9HEN8	Q9hfn8 candida rug
39	4	7.1	11	4	Q9UEL0	Q9uel0 homo sapien
40	4	7.1	11	10	Q9S8X4	Q9s8x4 glycine max
41	4	7.1	12	2	Q9R3B3	Q9r3b3 helicobacte
42	4	7.1	12	14	Q85631	Q85631 avian retro
43	4	7.1	13	2	O47693	O47693 escherichia
44	4	7.1	13	4	Q9UET3	Q9uet3 homo sapien
45	4	7.1	14	11	Q9Z0G5	Q9z0g5 mus musculu

#### ALIGNMENTS

RESULT 1

Q9UBN9 Q9UBN9 PRELIMINARY; PRT; 39 AA.

AC Q9UBN9; 01-MAY-2000 (TREMREL. 13, Created)

DT 01-MAY-2000 (TREMREL. 13, Last sequence update)

DT 01-MAR-2001 (TREMREL. 16, Last annotation update)

DE E6-AP UBIQUITIN-PROTEIN LIGASE (FRAGMENT).

GN UBE3A.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RX MEDLINE=9812641; PubMed=9465301;

RA Kishino T., Wagstaff J.;

RT "Genomic organization of the UBE3A/E6-AP gene and related

RT pseudogenes.";

RL Genomics 47:101-107(1998).

RN [2]

RP SEQUENCE FROM N.A.

RA Hennies H.C., Buerger J., Sperling K., Reis A.;

RT "Mutations in the E6-AP gene (UBE3A) in patients with Angelman

RT syndrome.";

RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF009341; AAC39580.1; -

DR EMBL; AJ001113; CAA04540.1; -

DR InterPro; IPR000569; -

DR PROSITE; PS0237; HECT; 1.

KW Ligase.

FT NON\_TER

FT 1

SQ SEQUENCE 39 AA; 4441 MW; 763722F374FA7193 CRC64;

Query Match 10.7%; Score 6; DB 4; Length 39;

Best Local Similarity 100.0%; Pred. No. 50;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SSKEKL 13  
|||||

RT reductase from the gram-negative bacterium Comamonas testosteroni.";  
RL Eur. J. Biochem. 241:744-749(1996).  
CC -!- FUNCTION: ALONG WITH THE 3 ALPHA-HYDROXYSTEROID DEHYDROGENASE AND  
CC 3-OXO-REDUCTASE ACTIVITIES TOWARDS A VARIETY OF CIS OR TRANS FUSED  
CC A/B RING STEROIDS. IT ALSO REDUCES SEVERAL XENOBIOTIC CARBONYL  
CC COMPOUNDS, INCLUDING A METYRAPON-BASED CLASS OF INSECTICIDES, TO  
CC THE RESPECTIVE ALCOHOL METABOLITES.  
CC -!- CATALYTIC ACTIVITY: ANDROSTERONE + NAD(P)(+) =  
CC 5-ALPHA-ANDROSTANE-3,17-DIONE + NAD(P)H.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- INDUCTION: BY STEROIDS.  
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
CC (SDR) FAMILY.  
DR InterPro; IPR002198; -.  
DR PROSITE; PS00061; ADH\_SHORT; PARTIAL.  
KW Oxidoreductase; NAD.  
FT DOMAIN 10 >30 INVOLVED IN COFACTOR BINDING  
FT NON\_TER 30 30 (BY SIMILARITY).  
SQ SEQUENCE 30 AA; 2829 MW; 065E9CF03F1C5A29 CRC64;

Query Match 7.1%; Score 4; DB 1; Length 30;  
Best Local Similarity 100.0%; Pred. No. 9.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 VLEA 45  
|  
|  
|  
Db 19 VLEA 22

Search completed: September 15, 2001, 12:54:26  
Job time: 110 sec

```

RESULT 12
CT31_LITCI
ID CT31_LITCI STANDARD; PRT; 24 AA.
AC P81851; P81852; P81853;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE CITROPIN 3.1.2 [CONTAINS: CITROPIN 3.1.1; CITROPIN 3.1.1].
OS Litoria citropa (Australian blue mountains tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Metazoa; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.
OX NCBI_TaxID=94770;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=99435977; PubMed=10504394;
RA Wegener K.L., Wabnitz P.A., Carver J.A., Bowie J.H., Chia B.C.S.,
RA Wallace J.C., Tyler M.J.;
RT "Host defence peptides from the skin glands of the Australian blue
RT mountains tree-frog Litoria citropa. Solution structure of the
RT antibacterial peptide citropin 1.1.";
RL Eur. J. Biochem. 265:627-637(1999).
CC -1- TISSUE SPECIFICITY: DORSAL AND SUBMENTAL SKIN GLANDS.
FT PEPTIDE 1 24 CITROPIN 3.1.2.
FT PEPTIDE 1 23 CITROPIN 3.1.1.
FT PEPTIDE 1 22 CITROPIN 3.1.1.
SQ SEQUENCE 24 AA; 2614 MW; C9001E295BD0E15D CRC64;

Query Match 7.1%; Score 4; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KEKL 13
Db 7 KEKL 10

RESULT 13
FEDG_AMYME
ID FEDG_AMYME STANDARD; PRT; 24 AA.
AC P80707;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE FORMATE ESTER DEHYDROGENASE, GAMMA CHAIN (EC 1.2.99.-) (FEDH)
DE (FRAGMENT).
OS Amycolatopsis methanolica.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Pseudonocardineae; Pseudonocardaceae;
OC Amycolatopsis.
OX NCBI_TaxID=1814;
RN [1]
RP SEQUENCE.
RX STRAIN=NCIB 11946;
RX MEDLINE=96140591; PubMed=8554333;
RA Kim S.W., Luykx D.M.A.M., de Vries S., Duine J.A.;
RT "A second molybdoprotein aldehyde dehydrogenase from Amycolatopsis
RT methanolica NCIB 11946.";
RL Arch. Biochem. Biophys. 325:1-7(1996).
CC -1- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA AND A GAMMA
CC CHAIN.
KW Oxidoreductase.
FT NON_TER 24 24
FT SEQUENCE 24 AA; 2746 MW; A93A8EA007D0FC6B CRC64;

Query Match 7.1%; Score 4; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 26 RTLL 29
Db 19 RTLL 22

RESULT 14
PA21_MICNI
ID PA21_MICNI STANDARD; PRT; 27 AA.
AC P21790;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PHOSPHOLIPASE A2 ISOZYME 1 (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE
DE 2-ACYLHYDROLASE) (FRAGMENT).
OS Micurus nigrocinctus (Central American coral snake) (Gargantilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Micurus.
OX NCBI_TaxID=8635;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RA Mochca-Morales J., Martin B.M., Zamudio F.Z., Possani L.D.;
RT "Isolation and characterization of three toxic phospholipases from
RT the venom of the coral snake Micurus nigrocinctus.";
RL Toxicon 28:616-617(1990).
CC -1- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
CC 2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC -1- FUNCTION: INHIBITS NEUROMUSCULAR TRANSMISSION BY BLOCKING
CC ACETYLCHOLINE RELEASE FROM THE NERVE TERMINI. ACT PRESYNAPTICALLY.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC PHOSPHOCHOLINE + A FATTY ACID ANION
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
DR PIR; A35948; A35948.
DR HSSP; P00598; 1POB.
DR InterPro; IPR001211; .
DR Pfam; PF00068; phoslip; 1.
DR PROSITE; PS00118; PA2_HIS; PARTIAL.
DR PROSITE; PS00119; PA2_ASP; PARTIAL.
KW Hydrolase; Lipid degradation; Calcium; Presynaptic neurotoxin; Venom;
KW Multigene family.
FT NON_TER 27 27
FT SEQUENCE 27 AA; 3314 MW; 38637ECA600F49A0 CRC64;

Query Match 7.1%; Score 4; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 SVLE 44
Db 16 SVLE 19

RESULT 15
DIDH_COMTE
ID DIDH_COMTE STANDARD; PRT; 30 AA.
AC P80702;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 3-ALPHA-HYDROXYSTEROID DEHYDROGENASE (EC 1.1.1.50) (3-ALPHA-HSD)
DE (HYDROXYPROSTAGLANDIN DEHYDROGENASE) (FRAGMENT).
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX NCBI_TaxID=285;
RN [1]
RP SEQUENCE.
RX STRAIN=ATCC 11996;
RX MEDLINE=97100200; PubMed=8944761;
RA Oppermann U.C.T., Maser E.;
RT "Characterization of a 3 alpha-hydroxysteroid dehydrogenase/carbonyl

```

RT control of proline biosynthesis.";  
 RL J. Gen. Microbiol. 137:509-517(1991).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D90351; BAA14363.1; -;  
 DR EMBL; X53086; CRA37253.1; -;  
 DR PIR; S11643; S11643.  
 DR PIR; C49753; C49753.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 20 AA; 2248 MW; 4DD777735276674 CRC64;

Query Match 7.1%; Score 4; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KEKL 13  
 Db 12 KEKL 15

RESULT 9  
 CR34\_LITCE STANDARD; PRT; 22 AA.  
 AC P56241;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CAERIN 3.4.  
 OS Litoria caerulea.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;  
 OC Litoria.  
 ON NCBI\_TaxID=30344;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Parotoid gland;  
 RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;  
 RT "Peptides from Australian frogs. The structures of the caerins from  
 RL Litoria caerulea.";  
 RL J. Chem. Res. 138:910-936(1993).  
 CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL  
 CC CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN  
 CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.  
 CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL  
 CC GLANDS.  
 CC -!- MASS SPECTROMETRY: MW=2452; METHOD=FAB.  
 KW Antibiotic; Amphibian skin; Amidation.  
 FT MOD\_RES 22 22  
 SQ SEQUENCE 22 AA; 2455 MW; 3AB40B2200D43663 CRC64;

Query Match 7.1%; Score 4; DB 1; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 7.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 IREK 56  
 Db 6 IREK 9

RESULT 10  
 MOTI\_CANFA STANDARD; PRT; 22 AA.  
 ID MOTI\_CANFA  
 AC P19863;

DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE MOTILIN.  
 GN MLN.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 ON NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Intestine;  
 RX MEDLINE=83195948; PubMed=6844663;  
 RA Poitras P., Reeve J.R. Jr., Hunkapiller M.W., Hood L.E., Walsh J.H.;  
 RT "Purification and characterization of canine intestinal motilin.";  
 RL Regul. Pept. 5:197-208(1983).  
 CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE REGULATION OF  
 CC INTERDIGESTIVE GASTROINTESTINAL MOTILITY AND INDIRECTLY CAUSES  
 CC RHYTHMIC CONTRACTION OF DUODENAL AND COLONIC SMOOTH MUSCLE.  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC -!- SIMILARITY: BELONGS TO THE MOTILIN FAMILY.  
 CC PIR: S00189; S00189.  
 DR PIR: A60313; A60313.  
 KW Hormone.  
 FT UNSURE 1 1  
 SQ SEQUENCE 22 AA; 2685 MW; 4BECB840ABE0639F CRC64;

Query Match 7.1%; Score 4; DB 1; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 7.2e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 IREK 56  
 Db 13 IREK 16

RESULT 11  
 FLAL\_SULSH STANDARD; PRT; 23 AA.  
 AC Q9UWG6;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE 31/33 KDA FLAGELLIN (FRAGMENT).  
 OS Sulfolobus shibatae.  
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
 ON NCBI\_TaxID=2286;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=B12;  
 RX MEDLINE=96146545; PubMed=8550530;  
 RA Fequn D.M., Bayley D.P., Kostyukova A.S., Thomas N.A., Jarrell K.F.;  
 RT "Isolation and characterization of flagella and flagellin proteins  
 RT from the Thermoacidophilic archaea Thermoplasma volcanium and  
 RT Sulfolobus shibatae.";  
 RL J. Bacteriol. 178:902-905(1996).  
 CC -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO  
 CC FORM THE FILAMENTS OF FLAGELLA.  
 CC -!- PTM: GLYCOSYLATED.  
 CC -!- SIMILARITY: BELONGS TO THE ARCHAEAL FLAGELLIN FAMILY.  
 KW Flagella; Glycoprotein.  
 FT NON\_TER 23 23  
 SQ SEQUENCE 23 AA; 2362 MW; BB372157B44DB3AC CRC64;

Query Match 7.1%; Score 4; DB 1; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 ASVL 43  
 Db 17 ASVL 20

15-DEC-1998 (Rel. 37, Last annotation update)  
 H(2)-FORMING N5,N10-METHYLENETETRAHYDROMETHANOPTERIN DEHYDROGENASE  
 (EC 1.12.99.-) (H2-DEPENDENT METHYLENE-HAMPT DEHYDROGENASE)  
 (FRAGMENT).  
 GN HMD.  
 OS Methanobacterium wolfei.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanothermobacter.  
 OX NCBI\_TaxID=145261;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92394151; PubMed=1521540;  
 RA Ziringibl C., van Dongen W., Schwoerer B., von Buehner R.,  
 Richter M., Klein A., Thauer R.K.;  
 RT "H2-forming methylenetetrahydromethanopterin dehydrogenase, a novel  
 type of hydrogenase without iron-sulfur clusters in methanogenic  
 archaea.";  
 RL Eur. J. Biochem. 208:511-520(1992).  
 CC -|- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROMETHANOPTERIN +  
 H(+) = 5,10-METHYLTETRAHYDROMETHANOPTERIN + H(2).  
 CC -|- COFACTOR: ZINC (POSSIBLE).  
 CC -|- PATHWAY: INVOLVED IN METHANOGENESIS.  
 CC -|- SUBUNIT: HOMODIMER.  
 KW Oxidoreductase; Methanogenesis; Zinc.  
 FT NON\_TER 1  
 SQ SEQUENCE 19 AA; 1911 MW; OC17E9D7BF1F97C9 CRC64;

Query Match 7.1%; Score 4; DB 1; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 DAAS 41  
 DB 14 DAAS 17

RESULT 6  
 ID PHSL\_DESBN STANDARD; PRT; 19 AA.  
 AC P13066;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE PERIPLASMIC [NIFES] HYDROGENASE LARGE SUBUNIT (EC 1.18.99.1) (NIFESE  
 DE HYDROGENYLASE LARGE CHAIN) (FRAGMENT).  
 OS Desulfotribrio baculatus (strain Norway 4).  
 OC Bacteria; Proteobacteria; delta subdivision; Desulfomicrobium.  
 OX NCBI\_TaxID=875;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=88106446; PubMed=3322275;  
 RA Prickril B.C., He S.H., Li C., Menon N., Choi E.S., Przybyla A.E.,  
 Dervartanian D.V., Peck H.D. Jr., Fauque G., le Gall J., Teixeira M.,  
 Moura I., Moura J.G., Patil D., Huynh B.H.;  
 RT "Identification of three classes of hydrogenase in the genus,  
 Desulfotribrio.";  
 RL Biochem. Biophys. Res. Commun. 149:369-377(1987).  
 CC -|- CATALYTIC ACTIVITY: 2 REDUCED FERREDOXIN + 2 H(+) = 2 OXIDIZED  
 CC FERREDOXIN + H(2).  
 CC -|- COFACTOR: ONE NICKEL ION, TWO 4FE-4S CLUSTERS AND ONE  
 CC SLENOCYSTEINE.  
 CC -|- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.  
 CC -|- SUBCELLULAR LOCATION: PERIPLASMIC.  
 CC -|- SIMILARITY: BELONGS TO THE [NIFE]/[NIFES] HYDROGENASE LARGE  
 CC SUBUNIT FAMILY.  
 PIR: H27480; H27480.  
 DR InterPro: IPR001501;  
 DR PROSITE: PS00507; N1\_HGENASE\_L1; PARTIAL.  
 DR PROSITE: PS00508; N1\_HGENASE\_L2; PARTIAL.  
 KW Oxidoreductase; Periplasmic; Nickel; Selenum; Selenocysteine.  
 FT NON\_TER 19  
 SQ SEQUENCE 19 AA; 1942 MW; 2BFCD2D360F00367 CRC64;

Query Match 7.1%; Score 4; DB 1; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 6.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKIS 4  
 DB 12 KKIS 15

RESULT 7  
 ID DEF6\_DERFA STANDARD; PRT; 20 AA.  
 AC P49276;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE MITE ALLERGEN DER F 6 (EC 3.4.21.-) (DER F VI) (DF5) (FRAGMENT).  
 GN DERF6.  
 OS Dermatophagoides farinae (House-dust mite).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcoptiformes; Astigmata; Analgoidea; Pyroglyphidae;  
 OC Dermatophagoides.  
 OX NCBI\_TaxID=6954;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93327207; PubMed=8334537;  
 RA Yasueda H., Mita H., Akiyama K., Shida T., Ando T., Sugiyama S.,  
 Yamakawa H.;  
 RT "Allergens from Dermatophagoides mites with chymotryptic activity.";  
 RL Clin. Exp. Allergy 23:384-390(1993).  
 CC -|- CATALYTIC ACTIVITY: SPECIFICITY SIMILAR TO CHYMOTRYPSIN.  
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR MEROPS: S01.187; .  
 DR InterPro: IPR001254; .  
 DR PROSITE: PS00134; TRYPSIN\_HIS; PARTIAL.  
 DR PROSITE: PS00135; TRYPSIN\_SER; PARTIAL.  
 KW Hydrolyase; Serine protease; Allergen.  
 FT UNSURE 1 2  
 FT NON\_TER 20 20  
 SQ SEQUENCE 20 AA; 2043 MW; 30B885F9BF3288C7 CRC64;

Query Match 7.1%; Score 4; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ISLL 6  
 DB 16 ISLL 19

RESULT 8  
 ID YPRB\_SERMA STANDARD; PRT; 20 AA.  
 AC P22581;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE HYPOTHETICAL PROTEIN IN PROB 5'REGION (FRAGMENT).  
 OS Serratia marcescens.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Serratia.  
 OX NCBI\_TaxID=615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=SR41;  
 FC MEDLINE=91237315; PubMed=1851803;  
 RA Omori K., Suzuki S., Inai Y., Komatsubara S.;  
 RT "Analysis of the Serratia marcescens proBA operon and feedback



Db 15 RERIK 19

## RESULT 2

UC15\_MAIZE  
ID UC15\_MAIZE STANDARD; PRT; 14 AA.  
AC P80621;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 245)  
DE (FRAGMENT).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;  
OC Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE.  
RC TISSUE-Coleoptile;  
RA Touzet P., Ricciardi F., Morin C., Damerval C., Huet J.-C.,  
RA Pernollet J.-C., Zivy M., de Vienne D.;  
RT "The maize two dimensional gel protein database: towards an integrated  
RT genome analysis program.";  
RL Theor. Appl. Genet. 93:997-1005(1996).  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 4.8, ITS MW IS: 35.7 KDA.  
DR Maize-2DPAGE; P80621; COLEOPTILE.  
DR MaizeDB; 123947; -.  
FT NON\_TER 1 1  
FT NON\_TER 14 14  
SQ SEQUENCE 14 AA; 1396 MW; C68949275F404CD2 CRC64;

## Query Match

Best Local Similarity 7.1%; Score 4; DB 1; Length 14;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 TLLP 30

Db 3 TLLP 6  
||||

## RESULT 3

UC27\_MAIZE  
ID UC27\_MAIZE STANDARD; PRT; 15 AA.  
AC P80633;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 688)  
DE (FRAGMENT).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;  
OC Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE.  
RC TISSUE-Coleoptile;  
RA Touzet P., Ricciardi F., Morin C., Damerval C., Huet J.-C.,  
RA Pernollet J.-C., Zivy M., de Vienne D.;  
RT "The maize two dimensional gel protein database: towards an integrated  
RT genome analysis program.";  
RL Theor. Appl. Genet. 93:997-1005(1996).  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
CC PROTEIN IS: 6.4, ITS MW IS: 48.4 KDA.  
CC -!- SIMILARITY: TO XENOPUS HISTONE-BINDING PROTEIN NL/N2 AND RABBIT  
CC AND HUMAN NUCLEAR AUTOANTIGENIC SPERM PROTEIN.  
DR Maize-2DPAGE; P80633; COLEOPTILE.  
DR MaizeDB; 123958; -.  
FT NON\_TER 1 1  
FT NON\_TER 15 15

SQ SEQUENCE 15 AA; 1853 MW; CA0E12A5DAED8DC7 CRC64;

## Query Match

Best Local Similarity 7.1%; Score 4; DB 1; Length 15;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 BQLR 26

Db 3 BQLR 6  
||||

## RESULT 4

HSTB\_ECOLI  
ID HSTB\_ECOLI STANDARD; PRT; 18 AA.  
AC P01560;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE HEAT-STABLE ENTEROTOXIN ST-2 (ST-B).  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=18D / SEROTYPE 0.42:K86:H37;  
RX MEDLINE=81264141; PubMed=7021541;  
RA Chan S.-K., Giannella R.A.;  
RA "Amino acid sequence of heat-stable enterotoxin produced by  
RT Escherichia coli pathogenic for man.";  
RL J. Biol. Chem. 256:7744-7746(1981).  
RN [2]  
RP DISULFIDE BONDS.  
RX MEDLINE=87191003; PubMed=3552731;  
RA Shimonishi Y., Hidaka Y., Koizumi M., Hane M., Aimoto S., Takeda T.,  
RA Miwatani Y., Takeda Y.;  
RT "Mode of disulfide bond formation of a heat-stable enterotoxin (STh)  
RT produced by a human strain of enterotoxigenic Escherichia coli.";  
RL FEBS Lett. 215:165-170(1987).  
CC -!- FUNCTION: TOXIN WHICH ACTIVATES THE PARTICULATE FORM OF GUANYLATE  
CC CYCLASE AND INCREASES CYCLIC GMP LEVELS WITHIN THE HOST  
CC INTESTINAL EPITHELIAL CELLS.  
CC -!- DISEASE: BOTH HEAT-STABLE AND HEAT-LABILE ENTEROTOXINS ARE  
CC PRODUCED BY PATHOGENIC STRAINS OF E. COLI AND AFFECT THE DIGESTIVE  
CC TRACT OF MAMMALS.  
CC -!- SIMILARITY: BELONGS TO THE HEAT-STABLE ENTEROTOXIN FAMILY.  
DR PIR: A01823; OHEC2.  
DR HSP; P01559; IETN.  
DR InterPro; IPR001489; -.  
DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
KW Toxin; Enterotoxin.  
FT DISULFID 5 10  
FT DISULFID 6 14  
FT DISULFID 9 17  
SQ SEQUENCE 18 AA; 1978 MW; D0C975F49D600650 CRC64;

## Query Match

Best Local Similarity 7.1%; Score 4; DB 1; Length 18;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YCCE 23

Db 4 YCCE 7  
||||

## RESULT 5

HMD\_METWO  
ID HMD\_METWO STANDARD; PRT; 19 AA.  
AC P32441;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:52:36 : Search time 10.44 Seconds  
(without alignments)  
183.746 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_189  
Perfect score: 56  
Sequence: 1 KKISLLHSSKEKLRRIKY.....NDAASVLEATVDYVKYIREK 56

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 93435 seqs, 34255486 residues  
Word size : 0

Total number of hits satisfying chosen parameters: 4168  
Minimum DB seq length: 0  
Maximum DB seq length: 56  
Post-processing: Listing first 45 summaries  
Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5	8.9	26	1 PUTA_KLEPN	P23725 klebsiella
2	4	7.1	14	1 UC15_MAIZE	P80621 zea mays (m
3	4	7.1	15	1 UC27_MAIZE	P80633 zea mays (m
4	4	7.1	18	1 HSTB_ECOLI	P01560 escherichia
5	4	7.1	19	1 HMD_METWO	P32441 methanobact
6	4	7.1	19	1 PHSL_DESBN	P13066 desulfovibr
7	4	7.1	20	1 DEF6_DERFA	P49276 dermatophag
8	4	7.1	20	1 YPRB_SERMA	P22581 serratia ma
9	4	7.1	22	1 CR34_LITCE	P56241 litorea cae
10	4	7.1	22	1 MOTI_CANFA	P19863 canis fami
11	4	7.1	23	1 FLAL_SULSH	Q9u966 sulfolobus
12	4	7.1	24	1 CT31_LITCI	P81851 litorea cit
13	4	7.1	24	1 FEDG_AYME	P80707 amycolatos
14	4	7.1	27	1 PA21_MICNI	P21790 micrurus ni
15	4	7.1	30	1 DIDH_COMTE	P80702 comamonas t
16	4	7.1	30	1 FSAM_ODOSI	P49487 odoncella s
17	4	7.1	31	1 SODC_STRHE	P81163 striga herm
18	4	7.1	31	1 Y822_BORBU	O51762 borrelia bu
19	4	7.1	32	1 YH17_HAEN	P44295 haemophilus
20	4	7.1	34	1 LEC2_CYTSE	P22971 cytisus ses
21	4	7.1	34	1 Y870_HAEN	P44065 haemophilus
22	4	7.1	35	1 KPPR_PINPS	P81664 pinus pinas
23	4	7.1	35	1 LEC1_CYTSE	P22970 cytisus ses
24	4	7.1	35	1 RL7_BUCAP	P41188 buchera ap
25	4	7.1	36	1 AMPL_PIG	P28839 sus scrofa
26	4	7.1	36	1 RL7_COXBU	O87902 coxiella bu
27	4	7.1	37	1 PSBM_PINTH	P41608 pinus thunb
28	4	7.1	37	1 FYI_CHICK	P29203 gallus gall
29	4	7.1	37	1 RK36_PEA	P07815 pisum sativ
30	4	7.1	37	1 TXOF_HADVE	P81599 hadronyche
31	4	7.1	42	1 GBG7_MOUSE	Q61016 mus musculu
32	4	7.1	42	1 V11_BPT7	P03779 bacterioph
33	4	7.1	43	1 ALA4_HORSE	P38031 equus cabal

34	4	7.1	43	1 PSBN_ZAMFU	Q9msrl zamia furfu
35	4	7.1	43	1 TVBY_HUMAN	O14604 homo sapien
36	4	7.1	44	1 RL34_BACST	P23376 bacillus st
37	4	7.1	44	1 RL34_BACSU	P05647 bacillus su
38	4	7.1	44	1 RL34_COXBU	P45647 coxiella bu
39	4	7.1	44	1 RL34_HAEIN	P44370 haemophilus
40	4	7.1	44	1 RL34_HELPY	P56056 helicobacte
41	4	7.1	44	1 RL34_PSEAE	P29436 pseudomonas
42	4	7.1	44	1 RL34_PSEPU	P16498 pseudomonas
43	4	7.1	44	1 YCX9_ODOSI	P49835 odontella s
44	4	7.1	45	1 REPA_STRPN	P13920 streptococc
45	4	7.1	45	1 RL34_BACHD	Q9rca3 bacillus ha

ALIGNMENTS

RESULT 1  
ID PUTA\_KLEPN STANDARD; PRT; 26 AA.  
AC P23725;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE BIFUNCTIONAL PUTA PROTEIN [INCLUDES: PROLINE DEHYDROGENASE  
DE (EC 1.5.9.8) (PROLINE OXIDASE); DELTA-1-PYRROLINE-5-CARBOXYLATE  
DE DEHYDROGENASE (EC 1.5.1.12) (P5C DEHYDROGENASE)] (FRAGMENT).  
GN PUTA.  
OS Klebsiella pneumoniae.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Klebsiella.  
OX NCBI\_TaxID=573;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91100369; PubMed=1987164;  
RA Chen L.M., Maloy S.;  
RT "Regulation of proline utilization in enteric bacteria: cloning and  
RT characterization of the Klebsiella put control region.";  
RL J. Bacteriol. 173:783-790(1991).  
CC -!- FUNCTION: OXIDIZES PROLINE TO GLUTAMATE FOR USE AS A CARBON AND  
CC NITROGEN SOURCE AND ALSO FUNCTION AS A TRANSCRIPTIONAL REPRESSOR  
CC OF THE PUT OPERON.  
CC -!- CATALYTIC ACTIVITY: L-PROLINE + ACCEPTOR + H(2)O = (S)-L-  
CC PYRROLINE-5-CARBOXYLATE + REDUCED ACCEPTOR.  
CC -!- COFACTOR: FAD.  
CC -!- PATHWAY: PROLINE UTILIZATION.  
CC -!- INDUCTION: BY PROLINE, AUTOREPRESSION AND CATABOLITE REPRESSION,  
CC AND IS POTENTIALLY NITROGEN CONTROLLED.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; M63160; AAA25139.1; -;  
CC InterPro; IPR002086; -;  
DR PROSITE; PS00070; ALDEHYDE\_DEHYDR\_CYS; PARTIAL.  
DR PROSITE; PS00687; ALDEHYDE\_DEHYDR\_GLU; PARTIAL.  
KW Multifunctional enzyme; Oxidoreductase; Flavoprotein; FAD; NAD;  
KW Transcription regulation; Repressor; DNA-binding; Proline metabolism.  
FT NON\_TER 26  
SQ SEQUENCE 26 AA: 2824 MW; BB332D0DE504CE19 CRC64;

Query Match 8.9%; Score 5; DB 1; Length 26;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 15 RERIK 19  
|||||

**THIS PAGE BLANK (USPTO)**



S61469  
 p83/100 protein - Borrelia afzelii (strain P1e and others) (fragment)  
 C:Species: Borrelia afzelii  
 A:Variety: strain P1e and others  
 C:Date: 19-Mar-1997 #sequence\_revision 13-Mar-1998 #text\_change 08-Oct-1999  
 C:Accession: S61469; S61470; S61471  
 R:Roessler, D.; Eifert, H.; Jauris-Heipke, S.; Lehnert, G.; Preac-Mursic, V.; Teepe, J.  
 Med. Microbiol. Immunol. 184, 23-32, 1995  
 A:Title: Molecular and immunological characterization of the p83/100 protein of various  
 A:Reference number: S61461; MUID:96149106  
 A:Accession: S61469  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-48 <ROE>  
 A:Cross-references: EMBL:X81529; NID:g928973; PIDN:CAA57248.1; PID:g928974; EMBL:X81530;  
 A:Experimental source: strain P1e; strain PK17; strain PGau  
 C:Keywords: surface antigen

Query Match 8.9%; Score 5; DB 2; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKER 12  
 |||||  
 Db 27 SSKER 31

## RESULT 14

S61472  
 p83/100 protein - Borrelia afzelii (strain PWudi) (fragment)  
 C:Species: Borrelia afzelii  
 A:Variety: strain PWudi  
 C:Date: 19-Mar-1997 #sequence\_revision 05-Dec-1998 #text\_change 08-Oct-1999  
 C:Accession: S61472; S72307  
 R:Roessler, D.; Eifert, H.; Jauris-Heipke, S.; Lehnert, G.; Preac-Mursic, V.; Teepe, J.  
 Med. Microbiol. Immunol. 184, 23-32, 1995  
 A:Title: Molecular and immunological characterization of the p83/100 protein of various  
 A:Reference number: S61461; MUID:96149106  
 A:Accession: S61472  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-48 <ROE>  
 A:Cross-references: EMBL:X81535  
 A:Experimental source: strain PWudi  
 R:Roessler, D.  
 submitted to the EMBL Data Library, September 1994  
 A:Reference number: S72307  
 A:Accession: S72307  
 A:Molecule type: DNA  
 A:Residues: 1-7, F', 8-48 <ROW>  
 A:Cross-references: EMBL:X81535; NID:g928977; PIDN:CAA57254.1; PID:g928978  
 A:Experimental source: strain PWudi  
 C:Keywords: surface antigen

Query Match 8.9%; Score 5; DB 2; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 SSKER 12  
 |||||  
 Db 27 SSKER 31

## RESULT 15

C69435  
 hypothetical protein AF1484 - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C:Accession: C69435  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch  
 A:Reference number: A69250; MUID:98049343  
 A:Accession: C69435  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-53 <KLE>  
 A:Cross-references: GB:AE001000; GB:AE000782; NID:g2689323; PIDN:AAB89773.1; PID:g264

Query Match 8.9%; Score 5; DB 2; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 YVKGR 35  
 |||||  
 Db 18 YVKGR 22

Search completed: September 15, 2001, 12:53:41  
 Job time: 110 sec

R; Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996  
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

S.  
 A:Reference number: S74322; MUID:97061201  
 A:Accession: S75759  
 A>Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-39 <KAN>  
 A:Cross-references: EMBL:D64003; GB:AB001339; NID:gl001200; PIDN:BAAL0494.1; PID:d101114  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 C:Genetics:  
 A:Start codon: GTG

Query Match 8.9%; Score 5; DB 2; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KISLL 6  
 Db 9 KISLL 13

RESULT 9

hypotheical protein VCA1064 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)  
 C:Species: *Vibrio cholerae*  
 C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

A:Accession: A82382  
 R; Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P. I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A:Reference number: A82035; MUID:20406833  
 A:Accession: A82382  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-40 <HEI>  
 A:Cross-references: GB:AE004432; GB:AE003853; NID:g9658509; PIDN:AAF96958.1; GSPDB:GN001  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VCA1064  
 A:Map position: 2

Query Match 8.9%; Score 5; DB 2; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 KGRKN 37  
 Db 35 KGRKN 39

RESULT 10

T07248  
 hypotheical protein 42a - *Chlorella vulgaris* chloroplast  
 C:Species: chloroplast *Chlorella vulgaris*  
 C>Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000  
 C:Accession: T07248  
 R;Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakas Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
 A>Title: Complete nucleotide sequence of the chloroplast genome from the green alga *Chlo*

A:Reference number: Z15985; MUID:97303241  
 A:Accession: T07248  
 A>Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-42 <WAK>  
 A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57895.1; PID:g2224411  
 C:Genetics:

A:Genome: chloroplast  
 C:Keywords: chloroplast

Query Match 8.9%; Score 5; DB 2; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 SKEKL 13  
 Db 19 SKEKL 23

RESULT 11

F81505  
 hypotheical protein CP1078 [imported] - *Chlamydomophila pneumoniae* (strain AR39)  
 C:Species: *Chlamydomophila pneumoniae*, *Chlamydia pneumoniae*  
 C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
 C:Accession: F81505  
 R; Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe Nucleic Acids Res. 28, 1397-1406, 2000  
 A>Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39

A:Reference number: A81500; MUID:20150255  
 A:Accession: F81505  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-43 <REA>  
 A:Cross-references: GB:AE002264; GB:AE002161; NID:g7189984; PIDN:AAF38850.1; PID:g718  
 A:Experimental source: strain AR39, HL cells  
 C:Genetics:  
 A:Gene: CP1078

Query Match 8.9%; Score 5; DB 2; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKISL 5  
 Db 7 KKISL 11

RESULT 12

T07311  
 hypotheical protein 47b - *Chlorella vulgaris* chloroplast  
 C:Species: chloroplast *Chlorella vulgaris*  
 C>Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000  
 C:Accession: T07311  
 R;Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Na Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
 A>Title: Complete nucleotide sequence of the chloroplast genome from the green alga C

A:Reference number: Z15985; MUID:97303241  
 A:Accession: T07311  
 A>Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-47 <WAK>  
 A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57959.1; PID:g2224475  
 C:Genetics:  
 A:Genome: chloroplast  
 C:Keywords: chloroplast

Query Match 8.9%; Score 5; DB 2; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSKEK 12  
 Db 3 SSKEK 7

RESULT 13

Db 13 SLHS 17

## RESULT 3

S10050

ribosomal protein L19.e - fission yeast (Schizosaccharomyces pombe) (fragment)

N:Alternate names: ribosomal protein SP-L15

C:Species: Schizosaccharomyces pombe

C&gt;Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 30-Sep-1993

C:Accession: S10050

R:Otaka, E.; Higo, K.I.; Itoh, T.

Mol. Gen. Genet. 191, 519-524, 1983

A:Title: Yeast ribosomal proteins: VII. Cytoplasmic ribosomal proteins from Schizosaccha

A:Reference number: S07293; MUID:84038947

A:Accession: S10050

A:Molecule type: protein

A:Residues: 1-29 &lt;OTA&gt;

C:Superfamily: rat ribosomal protein L19

C:Keywords: protein biosynthesis; ribosome

## Query Match

Best Local Similarity 8.9%; Score 5; DB 2; Length 29;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 AASVL 43

|||||

Db 10 AASVL 14

## RESULT 4

G82071

hypothetical protein VC2477 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C&gt;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: G82071

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: G82071

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-31 &lt;HEI&gt;

A:Cross-references: GB:AE004317; GB:AE003852; NID:g9657050; PIDN:AAF95619.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2477

A:Map position: 1

## Query Match

Best Local Similarity 8.9%; Score 5; DB 2; Length 31;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29

|||||

Db 16 LRTLL 20

## RESULT 5

I61697

myosin - human (fragment)

C:Species: Homo sapiens (man)

C&gt;Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 05-Nov-1999

C:Accession: I61697

R:Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.

Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994

A:Title: Identification and overlapping expression of multiple unconventional myosin gen

A:Reference number: A55758; MUID:94294418

A:Accession: I61697

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-31 &lt;RES&gt;

A:Cross-references: GB:L29146; NID:g457254; PIDN:AAA20909.1; PID:g531139

## Query Match

Best Local Similarity 8.9%; Score 5; DB 2; Length 31;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 VLEAT 46

|||||

Db 24 VLEAT 28

## RESULT 6

I46598

myosin - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C&gt;Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 05-Nov-1999

C:Accession: I46598

R:Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.

Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994

A:Title: Identification and overlapping expression of multiple unconventional myosin

A:Reference number: A55758; MUID:94294418

A:Accession: I46598

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-31 &lt;BEM&gt;

A:Cross-references: GB:L29133; NID:g457343; PIDN:AAA20918.1; PID:g531149

## Query Match

Best Local Similarity 8.9%; Score 5; DB 2; Length 31;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 VLEAT 46

|||||

Db 24 VLEAT 28

## RESULT 7

S27307

surface-array protein - Aeromonas hydrophila (fragment)

C:Species: Aeromonas hydrophila

C&gt;Date: 22-Nov-1993 #sequence\_revision 09-Mar-1996 #text\_change 07-May-1999

C:Accession: S27307; A33184

R:Kokka, R.P.; Vedros, N.A.; Janda, J.M.

J. Gen. Microbiol. 138, 1229-1236, 1992

A:Title: Immunochemical analysis and possible biological role of an Aeromonas: hydroph

A:Reference number: S27307; MUID:92407495

A:Accession: S27307

A&gt;Status: preliminary

A:Molecule type: protein

A:Residues: 1-35 &lt;KO2&gt;

## Query Match

Best Local Similarity 8.9%; Score 5; DB 2; Length 35;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 DAASV 42

|||||

Db 5 DAASV 9

## RESULT 8

S75759

hypothetical protein sgl0001 - Synecocystis sp. (strain PCC 6803)

C:Species: Synecocystis sp.

A:Variety: PCC 6803

C&gt;Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C:Accession: S75759

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:51:51 ; Search time 12.45 Seconds  
(without alignments)  
342.633 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_134\_189  
Perfect score: 56  
Sequence: 1 KKISLLHSSKEKLRRIKY.....NDAAASVLEATVDYVKYIREK 56

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 12943

Minimum DB seq length: 0  
Maximum DB seq length: 56

Post-processing: Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	10.7	42	2 T07291	hypothetical prote
2	5	8.9	28	2 T14905	hypothetical prote
3	5	8.9	29	2 S10050	ribosomal protein
4	5	8.9	31	2 G82071	hypothetical prote
5	5	8.9	31	2 I61697	myosin - human (fr
6	5	8.9	31	2 I46598	myosin - pig (frag
7	5	8.9	35	2 S27307	surface-array prot
8	5	8.9	39	2 S75759	hypothetical prote
9	5	8.9	40	2 A82382	hypothetical prote
10	5	8.9	42	2 T07248	hypothetical prote
11	5	8.9	43	2 F81505	hypothetical prote
12	5	8.9	47	2 T07311	hypothetical prote
13	5	8.9	48	2 S61469	p83/100 protein -
14	5	8.9	48	2 S61472	p83/100 protein -
15	5	8.9	53	2 G69435	hypothetical prote
16	5	8.9	53	2 G82813	hypothetical prote
17	5	8.9	54	2 A82652	hypothetical prote
18	4	7.1	6	2 B33932	Ig mu chain D regi
19	4	7.1	6	2 PT0560	T-cell receptor be
20	4	7.1	9	2 A45199	L-lysophosphorin -
21	4	7.1	12	2 C49215	urease (EC 3.5.1.5
22	4	7.1	14	2 PS0252	16K protein 5404 -
23	4	7.1	14	2 PC4382	dehydrin 4.5K poly
24	4	7.1	15	2 B49177	21K protein p2, mi
25	4	7.1	17	1 A05168	conotoxin G [vali
26	4	7.1	17	2 E53113	gingipain, 44K - p
27	4	7.1	17	2 A35550	adrenocortical cel
28	4	7.1	18	1 OHEC2	heat-stable entero
29	4	7.1	18	2 S74195	epoxide hydrolase

30	4	7.1	18	2 A60103	heat-stable entero
31	4	7.1	19	2 H27480	hydrogenase (EC 1.
32	4	7.1	19	2 B61056	cytochrome P450 PB
33	4	7.1	20	2 B37520	glutathione transf
34	4	7.1	20	2 D49215	urease (EC 3.5.1.5
35	4	7.1	20	2 I38417	HLA-A11 - human (f
36	4	7.1	20	2 S10680	probable 7-ethoxyc
37	4	7.1	20	2 C49753	hypothetical prote
38	4	7.1	20	2 S18582	hypothetical prote
39	4	7.1	20	2 T50757	puFK protein [impo
40	4	7.1	20	2 A60530	hycosporin - Japa
41	4	7.1	20	2 T03773	probable histone H
42	4	7.1	21	2 PT0431	leucyl aminopeptid
43	4	7.1	22	2 S00189	motilin - dog (ten
44	4	7.1	22	2 B33174	sormatin - sorghum
45	4	7.1	22	2 H49410	t-complex polypept

ALIGNMENTS

RESULT 1

T07291  
hypothetical protein 42c - Chlorella vulgaris chloroplast  
C:Species: Chloroplast Chlorella vulgaris  
C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000  
C:Accession: T07291  
R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Na  
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga C  
A:Reference number: Z15985; MUID:97303241  
A:Accession: T07291  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-42 <WAK>  
A:Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57939.1; PID:g2224455  
C:Genetics:  
A:Genome: chloroplast  
C:Keywords: chloroplast

Query Match 10.7%; Score 6; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKISLL 6  
Db 34 KKISLL 39

RESULT 2

T14905  
hypothetical protein - parsley  
C:Species: Petroselinum crispum (parsley)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T14905  
R:Feldbrugge, M.; Sprenger, M.; Dinkelbach, M.; Yazaki, K.; Harter, K.; Weisshaar, B.  
Plant Cell 6, 1607-1621, 1994  
A:Title: Functional analysis of a light-responsive plant bZIP transcriptional regulat  
A:Reference number: Z18259; MUID:95128172  
A:Accession: T14905  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-28 <FEL>  
A:Cross-references: EMBL:S75395; NID:g913201; PID:el94726

Query Match 8.9%; Score 5; DB 2; Length 28;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SLLHS 8  
|||||



**THIS PAGE BLANK (USPTO)**

---

Db 4 LRTLL 8  
|||||

RESULT 13  
PCT-US94-12985-13  
; Sequence 13, Application PC/TUS9412985  
; GENERAL INFORMATION:  
; APPLICANT: The Board of Trustees for the Leland Stanford Junior  
; APPLICANT: University  
; TITLE OF INVENTION: SURFACE MEMBRANE PROTEINS AND THEIR  
; TITLE OF INVENTION: EFFECT ON IMMUNE RESPONSE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER-READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/12985  
; FILING DATE: 10-NOV-1994  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/150,493  
; FILING DATE: 10-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rowland, Bertram I  
; REGISTRATION NUMBER: 20,015  
; REFERENCE/DOCKET NUMBER: FP-58976-PC/BIIR  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US94-12985-13

Query Match 8.9%; Score 5; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29  
|||||  
Db 4 LRTLL 8

RESULT 14  
US-08-433-613-9  
; Sequence 9, Application US/08433613A  
; Patent No. 6162434  
; GENERAL INFORMATION:  
; APPLICANT: Buelow, Roland  
; TITLE OF INVENTION: Cytomodulating Peptide for Inhibiting Lymphocyte  
; TITLE OF INVENTION: Activity  
; FILE REFERENCE: A61008/RFT/TAL  
; CURRENT APPLICATION NUMBER: US/08/433,613A  
; CURRENT FILING DATE: 1995-05-03  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9

; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-08-433-613-9

Query Match 8.9%; Score 5; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29  
|||||  
Db 8 LRTLL 12

RESULT 15  
US-08-433-613-48  
; Sequence 48, Application US/08433613A  
; Patent No. 6162434  
; GENERAL INFORMATION:  
; APPLICANT: Buelow, Roland  
; TITLE OF INVENTION: Cytomodulating Peptide for Inhibiting Lymphocyte  
; TITLE OF INVENTION: Activity  
; FILE REFERENCE: A61008/RFT/TAL  
; CURRENT APPLICATION NUMBER: US/08/433,613A  
; CURRENT FILING DATE: 1995-05-03  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-08-433-613-48

Query Match 8.9%; Score 5; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29  
|||||  
Db 8 LRTLL 12

Search completed: September 15, 2001, 12:53:19  
Job time: 104 sec

APPLICANT: CLAYBERGER, CAROL A.  
APPLICANT: KRENSKY, ALAN M.  
APPLICANT: PARHAM, PETER  
TITLE OF INVENTION: CYTOTOXIC T-CELL LYMPHOCYTE ("CTL")  
TITLE OF INVENTION: ACTIVITY REGULATION BY CLASS I MHC PEPTIDES  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 PENNSYLVANIA AVENUE, NW, STE 5500  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222.851  
FILING DATE: 05-APR-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: MILLMAN, ROBERT A.  
REGISTRATION NUMBER: 36,217  
REFERENCE/DOCKET NUMBER: 28600-20200.22  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 494-0792  
TELEX: 90-4030 MRSNFOERSWSH  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-222-851-14

Query Match 8.9%; Score 5; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29  
| | | | |  
Db 4 LRTLL 8

RESULT 11  
US-08-440-504A-9  
Sequence 9, Application US/08440504A  
Patent No. 5753625  
GENERAL INFORMATION:  
APPLICANT: Buelow, Roland  
TITLE OF INVENTION: Treatment for Inhibiting the Progression  
TITLE OF INVENTION: of Autoimmune Disease  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440.504A  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-60130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 949-8711  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-440-504A-9  
Query Match 8.9%; Score 5; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29  
| | | | |  
Db 4 LRTLL 8

RESULT 12  
US-08-440-504A-14  
Sequence 14, Application US/08440504A  
Patent No. 5753625  
GENERAL INFORMATION:  
APPLICANT: Buelow, Roland  
TITLE OF INVENTION: Treatment for Inhibiting the Progression  
TITLE OF INVENTION: of Autoimmune Disease  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440.504A  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-60130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 949-8711  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-440-504A-14

Query Match 8.9%; Score 5; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/12679  
FILING DATE: 30-DEC-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cranfill, Raymond B  
REGISTRATION NUMBER: 32,845  
REFERENCE/DOCKET NUMBER: RATH-10016PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-322-5333  
TELEFAX: 415-322-5499  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US93-12679-5

Query Match 8.9%; Score 5; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05; Indels 0;  
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

Qy 8 SSKK 12  
Db 2 SSKK 6

RESULT 8  
US-08-105-416-36  
; Sequence 36, Application US/08105416  
; Patent No. 5639958  
; GENERAL INFORMATION:  
; APPLICANT: Olsson, Lennart  
; APPLICANT: Goodenow, Robert S  
; APPLICANT: Goldstein, Avram  
; TITLE OF INVENTION: Class I MHC Modulation of Surface  
; TITLE OF INVENTION: Receptor Activity  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fierh, Hohbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/105,416  
; FILING DATE: 12-AUG-1993  
; CLASSIFICATION: 435  
; NAME: Rowland, Bertram I.  
; REGISTRATION NUMBER: 20015  
; REFERENCE/DOCKET NUMBER: A55115-4/BIR  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

US-08-105-416-36

Query Match 8.9%; Score 5; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05; Indels 0;  
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

Qy 25 LRTLL 29  
Db 1 LRTLL 5

RESULT 9  
US-08-473-656A-36  
; Sequence 36, Application US/08473656A  
; Patent No. 5853999  
; GENERAL INFORMATION:  
; APPLICANT: Olsson, Lennart  
; APPLICANT: Goodenow, Robert S  
; APPLICANT: Goldstein, Avram  
; TITLE OF INVENTION: Class I MHC Modulation of Surface  
; TITLE OF INVENTION: Receptor Activity  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fierh, Hohbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/473,656A  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/105,416  
; FILING DATE: 12-AUG-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rowland, Bertram I.  
; REGISTRATION NUMBER: 20015  
; REFERENCE/DOCKET NUMBER: A55115-4/BIR  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1989  
; TELEFAX: 415-398-3249  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-473-656A-36

Query Match 8.9%; Score 5; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05; Indels 0;  
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

Qy 25 LRTLL 29  
Db 1 LRTLL 5

RESULT 10  
US-08-222-851-14  
; Sequence 14, Application US/08222851  
; Patent No. 5723128  
; GENERAL INFORMATION:

;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
US-08-465-615-1

Query Match 10.7%; Score 6; DB 1; Length 27;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 LEATVD 48  
|||||  
Db 20 LEATVD 25

RESULT 5  
US-07-944-143C-17  
; Sequence 17, Application US/07944143C  
; Patent No. 5719064  
; GENERAL INFORMATION:  
; APPLICANT: Scofield, R. Hal  
; APPLICANT: Harley, John B.  
; TITLE OF INVENTION: Peptide Diagnostics and Therapeutics for  
; TITLE OF INVENTION: Spondyloarthropathies  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center  
; STREET: 1201 West Peachtree Street  
; CITY: Atlanta  
; STATE: GA  
; COUNTRY: USA  
; ZIP: 30309-3450

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/944,143C  
FILING DATE:  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)873-8794  
TELEFAX: (404)873-8795  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Salmonella typhimurium

Query Match 8.9%; Score 5; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LRTLL 29  
|||||  
Db 2 LRTLL 6

RESULT 6  
PCT-US93-08214-17  
; Sequence 17, Application PC/TUS9308214  
; GENERAL INFORMATION:  
; APPLICANT: Oklahoma Medical Research Foundation  
; TITLE OF INVENTION: Peptide Diagnostics and Therapeutics for Spondyloarthrop  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street, Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: United States  
; ZIP: 30309-4530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08214  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRF138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 815-6508  
TELEFAX: (404) 815-6555  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Salmonella typhimurium

Query Match 8.9%; Score 5; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LRTLL 29  
|||||  
Db 2 LRTLL 6

RESULT 7  
PCT-US93-12679-5  
; Sequence 5, Application PC/TUS9312679  
; GENERAL INFORMATION:  
; APPLICANT: Rath, Matthias  
; TITLE OF INVENTION: SYNTHETIC OLIGOPEPTIDES ANALOGOUS TO  
; TITLE OF INVENTION: PROTEIN SIGNAL SEQUENCES -METHODS OF IDENTIFICATION AND  
; TITLE OF INVENTION: METHODS OF USE  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SHELDON & MAX  
; STREET: 401 Florence Street  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

```
Db 30 SKEKLR 36
RESULT 2
; Sequence 8, Application US/08986837
; Patent No. 6221676
; GENERAL INFORMATION:
; APPLICANT: Lam, Bing K.
; APPLICANT: Penrose, John F.
; APPLICANT: Frank, Austen K.
; TITLE OF INVENTION: DNA ENCODING HUMAN LEUKOTRIENE C4 SYNTHASE,
; TITLE OF INVENTION: POLYPEPTIDES, AND USES THEREFOR
; FILE REFERENCE: Docket No. 6221676 0092662-0019 (BWH269)
; CURRENT APPLICATION NUMBER: US/08/986.837
; CURRENT FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 08/246.991
; EARLIER FILING DATE: 1994-05-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-986-837-8

Query Match 10.7%; Score 6; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LRTLPL 30
      |||||
Db 8 LRTLPL 13

RESULT 3
US-08-447-702-1
; Sequence 1, Application US/08447702
; Patent No. 5629190
; GENERAL INFORMATION:
; APPLICANT: Petre, Dominique
; APPLICANT: Cerbeleaud, Edith
; APPLICANT: Levy-Schil, Sophie
; APPLICANT: Crouzet, Joel
; TITLE OF INVENTION: Polypeptides Possessing A Nitrilase
; TITLE OF INVENTION: Activity, DNA Sequence Coding for Said Polypeptides,
; Patent No. 5629190
; TITLE OF INVENTION: Expression Cassettes and Host Microorganisms Enabling Them
; TITLE OF INVENTION: to be Obtained, and Method of Converting Nitriles to Carboxy-
; TITLE OF INVENTION: lates by Means of Said Polypeptides
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447.702
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/194,588
; FILING DATE: 10-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92-09-882

Query Match 10.7%; Score 6; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 LEATVD 48
      |||||
Db 20 LEATVD 25

RESULT 4
US-08-465-615-1
; Sequence 1, Application US/08465615
; Patent No. 5635391
; GENERAL INFORMATION:
; APPLICANT: PETRE, Dominique
; APPLICANT: CERBELEAUD, Edith
; APPLICANT: LEVY-SCHIL, Sophie
; APPLICANT: CROUZET, Joel
; TITLE OF INVENTION: POLYPEPTIDES POSSESSING A NITRILASE
; TITLE OF INVENTION: ACTIVITY, DNA SEQUENCE CODING FOR SAID POLYPEPTIDES,
; TITLE OF INVENTION: EXPRESSION CASSETTES AND HOST MICROORGANISMS ENABLING THEM
; TITLE OF INVENTION: TO BE OBTAINED, AND METHOD OF CONVERTING NITRILES TO
; TITLE OF INVENTION: CARBOXYLATES BY MEANS OF SAID POLYPEPTIDE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465.615
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/194,588
; FILING DATE: 10-FEB-1994
; APPLICATION NUMBER: FR 9209882
; FILING DATE: 10-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feuzy, Shaon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 003025-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: September 15, 2001, 12:51:35 ; Search time 12.43 Seconds  
(without alignments)  
92.764 Million cell updates/sec

Title: us-09-389-000-2\_copy\_134\_189  
Perfect score: 56  
Sequence: 1 KKTSLHSSKEKLRRIRIKY.....NDAASVLEATVDYVKYIREK 56

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141908

Minimum DB seq length: 0  
Maximum DB seq length: 56

Post-processing: Listing first 45 summaries

Database : Issued Patents\_AA.\*  
1: /cgnl\_7/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgnl\_7/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgnl\_7/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgnl\_7/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgnl\_7/ptodata/1/iaa/PCTUS.COMB.pep.\*  
6: /cgnl\_7/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	12.5	52	3	US-08-630-916A-76
2	6	10.7	15	4	US-08-986-837-8
3	6	10.7	27	1	US-08-447-702-1
4	6	10.7	27	1	US-08-465-615-1
5	5	8.9	6	1	US-07-944-143C-17
6	5	8.9	6	5	PCT-US93-08214-17
7	5	8.9	6	5	PCT-US93-12679-5
8	5	8.9	7	1	US-08-105-416-36
9	5	8.9	7	2	US-08-473-656A-36
10	5	8.9	10	1	US-08-222-851-14
11	5	8.9	10	1	US-08-440-504A-9
12	5	8.9	10	1	US-08-440-504A-14
13	5	8.9	10	5	PCT-US94-12985-13
14	5	8.9	14	4	US-08-433-613-9
15	5	8.9	14	4	US-08-433-613-48
16	5	8.9	15	2	US-08-750-856A-8
17	5	8.9	16	1	US-08-438-123-1
18	5	8.9	16	1	US-07-944-143C-16
19	5	8.9	16	5	PCT-US93-08214-16
20	5	8.9	17	1	US-07-976-872B-1
21	5	8.9	17	1	US-07-976-872B-2
22	5	8.9	17	1	US-07-976-872B-3
23	5	8.9	17	1	US-07-976-872B-4
24	5	8.9	17	1	US-08-105-416-12
25	5	8.9	17	1	US-08-105-416-13
26	5	8.9	17	1	US-08-105-416-18
27	5	8.9	17	1	US-08-105-416-19

28	5	8.9	17	1	US-08-105-416-20	Sequence 20, Appl
29	5	8.9	17	1	US-08-105-416-27	Sequence 27, Appl
30	5	8.9	17	1	US-08-105-416-29	Sequence 29, Appl
31	5	8.9	17	1	US-08-105-416-30	Sequence 30, Appl
32	5	8.9	17	1	US-08-105-416-37	Sequence 37, Appl
33	5	8.9	17	1	US-08-105-416-38	Sequence 38, Appl
34	5	8.9	17	2	US-08-473-656A-12	Sequence 12, Appl
35	5	8.9	17	2	US-08-473-656A-13	Sequence 13, Appl
36	5	8.9	17	2	US-08-473-656A-18	Sequence 18, Appl
37	5	8.9	17	2	US-08-473-656A-19	Sequence 19, Appl
38	5	8.9	17	2	US-08-473-656A-20	Sequence 20, Appl
39	5	8.9	17	2	US-08-473-656A-27	Sequence 27, Appl
40	5	8.9	17	2	US-08-473-656A-29	Sequence 29, Appl
41	5	8.9	17	2	US-08-473-656A-30	Sequence 30, Appl
42	5	8.9	17	2	US-08-473-656A-37	Sequence 37, Appl
43	5	8.9	17	2	US-08-473-656A-38	Sequence 38, Appl
44	5	8.9	17	3	US-08-483-931B-12	Sequence 12, Appl
45	5	8.9	17	3	US-08-483-931B-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1  
US-08-630-916A-76  
; Sequence 76, Application US/08630916A  
; Patent No. 6011137  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; APPLICANT: Kay, Brian K.  
; APPLICANT: Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/630,916A  
; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 896-8864/9741  
; INFORMATION FOR SEQ ID NO: 76:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 52 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-630-916A-76

Query Match 12.5%; Score 7; DB 3; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 SSKEKLRL 14  
|||||||



PT useful in pharmaceuticals, cosmetics, veterinary applications and  
PT agrochemicals  
XX  
PS Disclosure; Page 38; 55pp; English.  
XX  
CC The present invention relates to a method for identifying  
CC physico-chemical and/or topological parameters associated with biological  
CC activity. The method involves selecting the first subset from  
CC predetermined set of physico-chemical parameters, determining their value  
CC of function, and selecting the second subset from physico-chemical  
CC parameters based on the values of function, such that each second subset  
CC is more closely associated with the activity than the first subset. The  
CC selected physico-chemical parameters are useful for developing criteria  
CC for screening candidate molecules and are suitable for use in silico  
CC screening of compounds. The compounds may be used in pharmaceuticals,  
CC cosmetics, agrochemicals, biomaterials and in veterinary applications. It  
CC is also useful as an antibiotic or antifungal agent. The present  
CC sequence is an immunosuppressive peptide, Dk.75-84, identified by in  
CC silico screening. The immunosuppressive activity of the peptide that  
CC prevents allograft rejection is tested in a heterotopic allograft model  
XX of mouse.  
XX  
SQ Sequence 10 AA;

Query Match 8.9%; Score 5; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 25 LRTLL 29  
Db 4 lrtll 8

Search completed: September 15, 2001, 12:52:59  
Job time: 109 sec

XX PS Example 1; Page 19; 41pp; English.

XX CC Peptides AAW33784-98 and AAW33778-9 were assayed for their immunomodulating activity. A peptide-type compound or variant is claimed which has immunomodulating activity, including the N-terminal acylated and/or C-terminal amidated or esterified forms of up to 60 amino acids, where the peptide-type compound comprises the formula: A-B, where A, B = (R aa76-77L), (aa79-84) or (aa84-79) (Laa77-76R); aa76 = E or V; aa77 = D, S or N; aa79 = R or G; aa80 = I or N; aa81, aa84 = a hydrophobic or small amino acid; aa82 = R or L; aa83 = G or R; and aa represents amino acid. The sequence in the brackets may optionally be absent or truncated at any peptide type bond within the brackets. The compounds comprise amino acid sequences related to a Class I HLA-B alpha1 domain (positions 79-84). They can be used to inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in a host or in vitro. They can also be used in combination with antigenic peptides or proteins of interest to activate CTLs. They can also inhibit the proliferation of T cells in response to anti-CD3. The peptide can be used for preventing rejection of transplants or for treating autoimmune diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus. The products can also be used for detection and diagnosis.

XX SQ Sequence 10 AA;

Query Match 8.9%; Score 5; DB 19; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LRTLL 29  
| | | | |  
DB 4 lrtll 8

RESULT 14  
AAV50273  
ID AAV50273 standard; Peptide; 10 AA.  
XX AC AAV50273;  
XX DT 12-JAN-2000 (first entry)  
XX DE Neutrophil-activating pancreatic derived peptide 73.  
XX KW Cell activation; pancreas; treatment; cardiovascular disease; trauma; inflammatory disease; autoimmune diseases; arthritis; diabetes; stroke; organ rejection; ischemia; Alzheimer's disease; myocardial infarction; haemorrhagic shock; diabetic retinopathy; venous insufficiency; angina; trauma; protease inhibitor; hypertension; sepsis.  
XX OS Unidentified.  
XX PN WO9946367-A2.  
XX PD 16-SEP-1999.  
XX PF 11-MAR-1999; 99WO-US05247.  
XX PR 11-MAR-1998; 98US-0038894.  
XX PA (CELL-) CELL ACTIVATION INC.  
XX PA (BEGG ) UNIV CALIFORNIA.  
XX PA (SCRI ) SCRIPPS RES INST.  
XX PI Stoughton RB, Schmid-Schonbein GW, Hugli TE, Kistler E;  
XX DR WPI; 1999-580234/49.  
XX PT Use of cell activating compositions in developing products for diagnosis and treatment of e.g. cardiovascular, inflammatory, autoimmune or Alzheimer's disease, trauma, arthritis, organ rejection, diabetes, stroke or ischemia -

XX PS Example 9; Page 183; 184pp; English.

XX CC This invention describes a novel method for the use and preparation of cell activating compositions which involves preparing a cell activating composition comprising (a) homogenizing pancreatic tissue in buffer at about neutral or higher pH to produce a homogenate; (b) removing particulates from the homogenate; (c) optionally incubating the resulting homogenate, with particulates removed, with a protease; and (d) fractionating the homogenate and selecting fractions that exhibit cell activation activity. The methods can be used for improving treatment outcome or reducing risk of treatment of e.g. cardiovascular disease, inflammatory disease, trauma, autoimmune diseases, arthritis, organ rejection, diabetes and diabetic complications, stroke, ischemia, Alzheimer's disease, myocardial infarction, haemorrhagic shock, diabetic retinopathy, diabetes, venous insufficiency, unstable angina or trauma. They can be used in the veterinary treatment of a non-human subject. Protease inhibitors can be used to lower cell activation resulting from these diseases and deficiencies. The detection of an elevated level of hydrogen peroxide can be used to detect an inflammatory condition. An elevated level of hydrogen peroxide in plasma or whole blood and in the presence of superoxide dismutase (SOD) indicates leukocyte up regulation, e.g. indicative of the onset of an acute cardiovascular disorders, such as disease onset or ischemic complications. An elevated level of hydrogen peroxide in plasma or whole blood and a low level in the presence of SOD is indicative of a chronic or immune compromised condition e.g. hypertension or sepsis. AAV50201-Y50334 represent peptides used in the method of the invention.

XX SQ Sequence 10 AA;

Query Match 8.9%; Score 5; DB 20; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 SVLEA 45  
| | | | |  
DB 1 svlea 5

RESULT 15  
AAV72484  
ID AAV72484 standard; peptide; 10 AA.  
XX AC AAV72484;  
XX DT 24-APR-2001 (first entry)  
XX DE Immunosuppressive peptide, Dk.75-84, to prevent allograft rejection.  
XX KW Immunosuppressive; allograft rejection; topological parameter; physico-chemical parameter; in silico screening; pharmaceutical; cosmetic; agrochemical; biomaterial; veterinary application.  
XX OS Unidentified.  
XX PN WO200079263-A2.  
XX PD 28-DEC-2000.  
XX PF 15-MAY-2000; 2000WO-EP04338.  
XX PR 18-JUN-1999; 99EP-0401526.  
XX PA (SYNT-) SYNT:EM SA.  
XX PI Lahana R, Clair P, Yasri A;  
XX DR WPI; 2001-091623/10.  
XX PT Identifying active candidate molecules on the basis of selected physico-chemical parameters, for in silico screening of compounds

PF 05-APR-1996; 96WO-US04710.  
 PR 12-MAY-1995; 95US-0440504.  
 XX (SANG-) SANGSTAT MEDICAL CORP.  
 PA Buelow R;  
 PI WPI; 1996-518410/51.  
 XX Treatment of auto-immune disease by admin. of peptide(s) corresp. to  
 XX major histocompatibility complex antigens - esp. for delaying onset  
 XX of clinical symptoms of insulin dependent diabetes by modulating T  
 XX cell mediated attack on target cells  
 XX Example 1; Page 12; 24pp; English.  
 PS AAW07521-W07524, and AAW07527 represent T-cell modulating peptides that  
 XX can be used in the method of the invention. These sequences are based on  
 XX a portion of the generic peptide corresponding to residues 70-91 of the  
 XX alpha1-domain of the major histocompatibility complex (MHC) class I  
 XX antigen (see AAW07510). The method is for affecting the course of an  
 XX autoimmune disease involving T-cell mediated destruction of tissue in  
 XX mammals. These peptides are used especially to treat insulin-dependent  
 XX diabetes mellitus, preferably being administered during the pre-clinical  
 XX stage to delay onset of the disease. Other diseases that can be treated  
 XX are multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus  
 XX vulgaris, Sjogren's disease, thyroid disease, Hashimoto's thyroiditis,  
 XX myasthenia gravis, etc. The peptides modulate T-cell mediated attack on  
 XX autologous target cells, and may also reduce inflammation, swelling, and  
 XX release of cytokines, perforins, granzymes etc. associated with T cell  
 XX activation.  
 SQ Sequence 10 AA;

Query Match 8.9%; Score 5; DB 17; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LRTLL 29  
 Db 4 Lrtll 8  
 |||||

RESULT 12  
 AAW07517  
 ID AAW07517 standard; peptide; 10 AA.  
 AC AAW07517;  
 XX 04-AUG-1997 (first entry)  
 DT T-cell modulating peptide #6.  
 XX T-cell modulator; autoimmune disease; tissue destruction; alpha1-domain;  
 KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;  
 KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;  
 KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;  
 KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;  
 KW autologous target cell; cytokine release; T cell activation; therapy.  
 XX Synthetic.  
 OS WO9635443-A1.  
 XX 14-NOV-1996.  
 PD 05-APR-1996; 96WO-US04710.  
 XX 12-MAY-1995; 95US-0440504.  
 PR (SANG-) SANGSTAT MEDICAL CORP.  
 PA

XX Buelow R;  
 PI WPI; 1996-518410/51.  
 XX Treatment of auto-immune disease by admin. of peptide(s) corresp. to  
 XX major histocompatibility complex antigens - esp. for delaying onset  
 XX of clinical symptoms of insulin dependent diabetes by modulating T  
 XX cell mediated attack on target cells  
 XX Claim 7; Page 20; 24pp; English.  
 PS AAW07512-W07518 represent T-cell modulating peptides that can be used in  
 XX the method of the invention. These sequences are based on a portion of  
 XX the generic peptide corresponding to residues 70-91 of the alpha1-domain  
 XX of the major histocompatibility complex (MHC) class I antigen (see  
 XX AAW07510). The method is for affecting the course of an autoimmune  
 XX disease involving T-cell mediated destruction of tissue in mammals.  
 XX These peptides are used especially to treat insulin-dependent diabetes  
 XX mellitus, preferably being administered during the pre-clinical stage to  
 XX delay onset of the disease. Other diseases that can be treated are  
 XX multiple sclerosis, rheumatoid arthritis, psoriasis, pemphigus vulgaris,  
 XX Sjogren's disease, thyroid disease, Hashimoto's thyroiditis, myasthenia  
 XX gravis, etc. The peptides modulate T-cell mediated attack on autologous  
 XX target cells, and may also reduce inflammation, swelling, and release of  
 XX cytokines, perforins, granzymes etc. associated with T cell activation.  
 SQ Sequence 10 AA;

Query Match 8.9%; Score 5; DB 17; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LRTLL 29  
 Db 4 Lrtll 8  
 |||||

RESULT 13  
 AAW33785  
 ID AAW33785 standard; peptide; 10 AA.  
 XX AAW33785;  
 AC 19-JUN-1998 (first entry)  
 DT Peptide B2705.75-84 tested for immunomodulating activity.  
 XX Immunomodulating dimer; immunosuppressant drug; CTL activation;  
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;  
 KW rejection.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX WO9744351-A1.  
 XX 27-NOV-1997.  
 PD 22-MAY-1997; 97WO-US08689.  
 XX 24-MAY-1996; 96US-0653294.  
 PR (STRD ) UNIV LELAND STANFORD JUNIOR.  
 PA Buelow R, Clayberger C, Krensky AM;  
 PI WPI; 1998-086530/08.  
 XX New immunomodulating dimer peptide(s) - based on a Class I HLA-B  
 XX alpha-1 domain, used for preventing rejection of transplants or  
 XX treating autoimmune diseases

DT 12-NOV-1996 (first entry)  
 DE HLA-B2705.75-84.  
 XX  
 KW HLA; p74; alpha-helix; human-leucocyte-associated antigen; inhibitor;  
 KW T-cell lysate; membrane protein; mammal; heat shock protein; Hsc70; APC;  
 KW B cell; calcium influx; cytotoxic T lymphocyte; CTL; differentiation;  
 KW cytolysis; antigen presenting cell.  
 XX  
 OS Synthetic.  
 XX  
 OS WO9513288-A1.  
 PN 18-MAY-1995.  
 XX  
 PD 10-NOV-1994; 94WO-US12985.  
 XX  
 PF 10-NOV-1993; 93US-0150493.  
 XX  
 PR (STRD ) UNIV LELAND STANFORD JUNIOR.  
 XX  
 PA Clayberger C, Krensky AM;  
 XX  
 PI WPI; 1995-194027/25.  
 XX  
 DR Compsns. comprising lymphoid surface membrane proteins - which may  
 XX inhibit cytolytic activity and differentiation of CTLs.  
 XX  
 PS Example; Page 11; 29pp; English.  
 XX  
 CC AAR95413, and AAR95415-R95431 represent palindromes and fragments of  
 CC human-leucocyte-associated antigens. This sequence represents the  
 CC HLA-B2705.75-84. These sequences can be used to isolate the protein p74  
 CC from a T-cell lysate. p74 is a T-cell surface membrane protein  
 CC associated with T-cell activation in mammalian T-cells, and is also  
 CC immunologically cross reactive with the heat shock protein Hsc70. p74  
 CC is found in a limited number of cell types, but is particularly expressed  
 CC on B and T cells. p74 can be isolated by lysis of a suitable cell with  
 CC an amphoteric detergent, and then passed through an affinity column  
 CC containing a covalently bound HLA-B2702 palindromic peptide.  
 CC Compositions comprising the extracellular fragment of p74 combined with  
 CC HLA-B2702.60-84 (see AAR95416), induces calcium influx, and inhibits  
 CC cytotoxic T lymphocyte (CTL) differentiation or cytolysis. Candidate  
 CC compounds can be screened for their effect on the cytolytic activity of  
 CC T-cells, by combining them with the extracellular portion of p74 and  
 CC determining the amount of binding between the candidate compound and  
 CC p74. Modulation of CTL activity can be inhibited in a cellular  
 CC composition containing T-cells and antigen presenting cells (APCs), by  
 CC adding to the mix the extracellular portion of p74, in an amount  
 CC sufficient to compete with p74 for the binding of the p74 ligand.  
 XX  
 SQ Sequence 10 AA;  
 Query Match 8.9%; Score 5; DB 16; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 25 LRTLL 29  
 Db 4 lrtll 8  
 RESULT 10  
 AAR83075  
 ID AAR83075 standard; peptide; 10 AA.  
 AC AAR83075;  
 XX  
 XX 16-MAY-1996 (first entry)  
 DT HLA-B2702 CTL modulating peptide (B2702.75-84).  
 DE  
 XX

Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 immunosuppressant; graft versus host disorder; transplantation; therapy;  
 Class I MHC; HLA-B2702.  
 Synthetic.  
 WO9526979-A1.  
 12-OCT-1995.  
 05-APR-1995; 95WO-US04349.  
 05-APR-1994; 94US-0222851.  
 (STRD ) UNIV LELAND STANFORD JUNIOR.  
 Clayberger C, Krensky AM, Parham P;  
 WPI; 1995-358582/46.  
 Extension of acceptance period of transplants from MHC unmatched  
 donor hosts - using Class I B75-84 MHC antigen of the recipient  
 host  
 Example 14; Page 34; 80pp; English.  
 AAR83061-R83085, AAR83090-R83096 and AAR92907-R92913 represent fragments  
 of class I major histocompatibility complex (MHC) antigens. This  
 sequence corresponds to residues 75-84 of the alpha-1 domain of the  
 class I MHC HLA-B2702. These sequences can be used to extend the period  
 of acceptance by a recipient of a transplant from an MHC unmatched  
 donor. The peptides are administered to a patient in conjunction with a  
 subtherapeutic amount of an immunosuppressant. This is administered to  
 the patient for a limited period of time (compared to the lifetime  
 administration for current treatments). The peptides particularly  
 modulate (or inhibit) the activity of the cytotoxic T lymphocytes (CTLs)  
 of the patient.  
 Sequence 10 AA;  
 Query Match 8.9%; Score 5; DB 16; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 25 LRTLL 29  
 Db 4 lrtll 8  
 RESULT 11  
 AAW07524  
 ID AAW07524 standard; peptide; 10 AA.  
 AC AAW07524;  
 XX  
 XX 04-AUG-1997 (first entry)  
 DT T-cell modulating peptide Dk.  
 DE  
 XX  
 XX T-cell modulator; autoimmune disease; tissue destruction; alpha1-domain;  
 KW mammal; major histocompatibility complex; MHC class I; antigen; perforin;  
 KW insulin-dependent diabetes mellitus; multiple sclerosis; inflammation;  
 KW rheumatoid arthritis; psoriasis; pemphigus vulgaris; Sjogren's disease;  
 KW thyroid disease; Hashimoto's thyroiditis; myasthenia gravis; granzyme;  
 KW autologous target cell; cytokine release; T cell activation; therapy.  
 XX  
 OS Synthetic.  
 XX  
 XX WO9635443-A1.  
 XX  
 XX 14-NOV-1996.  
 XX

XX PS Example 26; Page 27; 47pp; English.

CC The patent discloses dynorphin-like polypeptides which are shorter

CC than dynorphin (7 to 9 amino acids) and which contain either a D-

CC amino acid and/or an N-alkyl deriv. of an L- or D-amino acid to

CC provide in-vivo stability upon intravenous administration. The

CC peptides have a much greater analgesic effect than dynorphin

CC due to their greater stability in the blood (e.g. IC50 values can be

CC as low as 0.04 nM when tested by the rabbit vas deferens method; cf.

CC 17.4 nM for dynorphin).

CC The present peptide is a specific example of the peptides.

XX SQ Sequence 9 AA;

Query Match 8.9%; Score 5; DB 15; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LRRER 17

Db 5 Lrrer 9

RESULT 7

AAR84869

ID AAR84869 standard; peptide; 9 AA.

AC AAR84869;

XX 30-MAY-1996 (first entry)

DT H2Kb 75-83 immunogenic peptide.

DE Antigen; epitope; cell mediated; immune specific; cancer;

XX infection; infestation; mucin-1; MUC-1; tumour; H2Kb 75-83;

KW immunogenic peptide.

XX Synthetic.

XX WO9527505-A1.

PN 19-OCT-1995.

XX 12-APR-1995; 95WO-US04540.

PF 12-APR-1994; 94US-0229606.

XX (BIOM-) BIOMIRA INC.

PA Ding L, Koganty RR, Longenecker BM, Reddish MA;

PI WPI; 1995-373528/48.

DR New cell-mediated immune-specific immunogenic compsns. - used in

XX prophylaxis and treatment of cancer, microbial infections, viral

PT infections and parasitic infestations

PS Example 1; Page 62; 141pp; English.

XX A claimed cell-mediated immune (CMI)-specific immunogenic compsn.

CC comprises a conjugate of a primary antigen bearing a primary epitope,

CC with an immunomodulatory peptide (IP), i.e. the present peptide.

CC The IP comprises an alloptide moiety of at least 5 amino acids,

CC whose sequence corresponds to a polymorphic region of a MHC

CC encoded polymorphic Class I or II antigen. The compsn. can be

CC used to elicit a CMI-specific response which is prophylactic, or

CC therapeutic for, e.g. microbial and viral infections, parasitic

CC infestations and cancer, partic. MUC-1 expressing tumour cells

CC when the present peptide is the IP, and a MUC-1 epitope is the

CC primary epitope.

XX

SQ Sequence 9 AA;

Query Match 8.9%; Score 5; DB 16; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29

Db 4 lrtll 8

RESULT 8

AAR41212

ID AAR41212 standard; peptide; 10 AA.

XX AAR41212;

XX 15-MAR-1994 (first entry)

DT Peptide fragment of Class I HLA peptide.

DE Human leukocyte antigen; HLA; peptide; transplantation; neoplasia;

XX parasitic disease; cytotoxic T lymphocyte; modulation.

KW Synthetic.

XX WO9317699-A.

PN 16-SEP-1993.

PD 25-FEB-1993; 93WO-US01758.

XX 02-MAR-1992; 92US-0844716.

PR (STRD ) UNIV LELAND STANFORD JUNIOR.

XX Clayberger CA, Krensky AM;

PI WPI; 1993-303134/38.

DR New peptide(s) based on Class I HLA antigen domains - used for

XX modulating cytotoxic T-lymphocyte activity towards targets

PS Claim 11; Page 54; 6lpp; English.

XX The peptide is used to modulate cytotoxic T-lymphocyte (CTL)

CC activity, either by inhibition or stimulation. It can be used

CC for inhibiting CTL toxicity in transplantations, for inducing CTL

CC activity in parasitic diseases and neoplasia and in studies on viral

CC infection. The peptide can also be used for identifying CTLs which

CC bind to it and removing subsets of CTLs from a T-cell composition.

CC This peptide sequence is more commonly found within larger peptide

CC compounds of not more than 30 amino acids in length.

XX SQ Sequence 10 AA;

Query Match 8.9%; Score 5; DB 14; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29

Db 4 lrtll 8

RESULT 9

AAR95423

ID AAR95423 standard; peptide; 10 AA.

XX AAR95423;

XX

```

XX PF 30-DEC-1993; 93WO-US12679.
XX PR 30-DEC-1992; 92US-0997727.
XX PA (RATH/) RATH M.
XX PI Rath M;
XX DR WPI; 1994-249399/30.
XX PT Identifying peptide signal sequences in a protein - and use of
XX PT their synthetic analogues for treating or preventing, e.g.
XX PT cardiovascular and auto-immune disease, infections and cancer.
XX PS Claim 18; Page 10; 28pp; English.
XX CC The sequence is that of a peptide signal sequence which can be used
XX CC to treat E. coli infections.
XX CC See also AAR59944-83.
XX SQ Sequence 6 AA;

Query Match 8.9%; Score 5; DB 15; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SSKK 12
Db 2 sskek 6

RESULT 6
AAR61019
ID AAR61019 standard; peptide; 9 AA.
XX AC AAR61019;
XX DT 23-APR-1995 (first entry)
XX DE Dynorphin-like polypeptide.
XX KW Dynorphin; opioid analgesic; stable; stability; intravenous.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "Metyr"
XX FT Misc-difference 8 /note= "D-Glu"
XX FT Modified-site 9 /note= "Arg-NH2"
XX PN EP614913-A.
XX XX
XX PD 14-SEP-1994.
XX PF 08-NOV-1985; 94EP-0107769.
XX PR 09-NOV-1984; 84JP-0236076.
XX XX
XX PA (EISA ) EISAI CO LTD.
XX PI Arakawa Y, Araki S, Ikeda M, Kaneko T, Nakazawa T;
XX PI Tachibana S, Tsuchiya Y, Yamatsu K, Yoshino H;
XX XX
XX DR WPI; 1994-281157/35.
XX PT New dynorphin polypeptide derivs. for use as analgesics - contg.
XX PT a D-amino acid or an N-alkyl deriv. of a D- or L-amino acid for
XX PT in vivo stability.

RESULT 4
AAR50267
ID AAR50267 standard; peptide; 6 AA.
XX AC AAR50267;
XX DT 13-OCT-1994 (first entry)
XX DE Peptide corresponding to a shared sequence of an HLA molecule.
XX KW Peptide; HLA; human leucocyte antigen; diagnosis; treatment;
XX KW autoimmune disease; uveitis; spondylitis; psoriasis;
XX KW inflammatory bowel disease; enteric bacteria;
XX KW Salmonella typhimurium.
XX OS Salmonella typhimurium.
XX PN W09405303-A.
XX XX
XX PD 17-MAR-1994.
XX PF 31-AUG-1993; 93WO-US08214.
XX PR 31-AUG-1992; 92US-0944143.
XX PA (OKLA ) UNIV OKLAHOMA STATE.
XX PI Harley JB, Scofield RH;
XX DR WPI; 1994-100843/12.
XX PT Peptide corresponding to shared sequences of HLA molecules and
XX PT enteric bacteria - used for the diagnosis and treatment of
XX PT auto-immune disorders, partic. spondylarthropathies.
XX PS Disclosure; Page 15; 58pp; English.
XX CC The peptide is a fragment of the HLA B27 hypervariable region
XX CC described in AAR50266. The peptide can be used for the diagnosis and
XX CC treatment of autoimmune disorders, e.g. spondyloarthropathies including
XX CC uveitis and spondylitis associated with inflammatory bowel disease
XX CC or psoriasis.
XX SQ Sequence 6 AA;

Query Match 8.9%; Score 5; DB 15; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 LRTLL 29
Db 2 lrtll 6

RESULT 5
AAR59948
ID AAR59948 standard; peptide; 6 AA.
XX AC AAR59948;
XX DT 14-FEB-1995 (first entry)
XX DE Peptide signal sequence for treating E. coli infections.
XX KW Therapeutic; metabolic-interactions; PSS; analogues.
XX OS Synthetic.
XX PN W09416328-A.
XX PD 21-JUL-1994.
```

PT Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -

XX Example 5; Page 36; 62pp; English.

XX The present sequence is that of a peptide derived from human  
 CC PHELIIX (see AAY79269), a novel transcription factor that is  
 CC normally expressed only in testis tissue, but which is up-regulated  
 CC in prostate and some other cancers. The peptide was conjugated to  
 CC keyhole limpet haemocyanin and used to raise polyclonal antiserum  
 CC in rabbit. The antiserum demonstrated specificity for PHELIIX and  
 CC may therefore be useful for assessing the expression of PHELIIX in  
 CC patient samples.

XX Sequence 15 AA;

Query Match 25.0%; Score 14; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 HSKSKLRRRIKY 20  
 Db 1 hsksklrreriky 14  
 |||||

RESULT 2

AAAB15385  
 ID AAB15385 standard; Protein; 27 AA.

XX AC AAB15385;

XX DT 06-DEC-2000 (first entry)

XX DE Comamonas testosteroni nitrilase protein N-terminus.

XX KW Nitrilase; hydrolysis; dinitrile; carboxylic acid; adiponitrile;  
 KW adipic acid; nylon 6,6; 5-cyanovaleic acid.

XX OS Comamonas testosteroni.

XX PN FR2694571-A1.

XX PD 11-FEB-1994.

XX PF 10-AUG-1992; 92FR-0009882.

XX PR 10-AUG-1992; 92FR-0009882.

XX PA (RHON ) RHONE POULENC CHIM.

XX PI Petre D, Cerebelaud E, Levy-Schil S, Crouzet J;

XX DR WPI; 1994-076687/10.

XX New gene for nitrilase from Comamonas testosteroni - esp. for  
 PT converting di-nitrile(s) to di-carboxylic acids, also the new enzyme  
 PT and transformed microorganisms

XX Example 1; Page 10; 33pp; French.

XX This sequence represents the N-terminal sequence of a nitrilase, able  
 CC to hydrolyse nitriles to carboxylic acids, isolated from the bacterium  
 CC Comamonas testosteroni. The N-terminal sequence was obtained by peptide  
 CC sequencing of the purified enzyme. The nitrilase, or microorganisms  
 CC which produce it, is useful for converting dinitriles NC.R-CN (R=1-10C  
 CC alkylene) to corresponding acids, e.g. adiponitrile to adipic acid (for  
 CC nylon 6,6 manufacture) or 5-cyanovaleic acid (as their NH4 salts). The  
 CC protein provides rapid and complete hydrolysis of dinitriles; contrast  
 CC known nitrilases with which hydrolysis of the second CN is usually very  
 CC slow.

XX

SQ Sequence 27 AA;

Query Match 10.7%; Score 6; DB 15; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 LEATVD 48  
 Db 20 leatvd 25  
 |||||

RESULT 3

AAP30295  
 ID AAP30295 standard; Protein; 6 AA.

XX AC AAP30295;

XX DT 20-APR-1992 (first entry)

XX DE Sequence of AAs 26-31 of the E. coli heat stable toxin  
 DE which determine an H-epitope.

XX KW Synthetic vaccine; antigen; allergen; immunological response;  
 KW antibody.

XX OS Escherichia coli.

XX PN EP93851-A.

XX PD 16-NOV-1983.

XX PF 11-MAR-1983; 83EP-0102392.

XX PR 15-MAR-1982; 82US-0358150.

XX PR 28-JAN-1983; 83US-0461802.

XX PR 12-JUN-1981; 81US-0272855.

XX PR 09-JAN-1981; 81US-0223558.

XX PR 16-DEC-1986; 86US-0942562.

XX PA (NYBL-) NEW YORK BLOOD CENT.

XX PI Hopp TP;

XX DR WPI; 1983-822049/47.

XX Synthetic vaccine - contains peptide residue coupled to higher  
 PT alkyl or alkenyl Gps. and with 6 amino acids in residue

XX Claim 50; Page 46; 54pp; English.

XX The inventors claim a synthetic vaccine which comprises a peptide  
 CC residue coupled to an alkyl or alkenyl gp. having at least 12C or  
 CC other lipophilic substance. The residue contains a sequence of 6 AAs  
 CC corresp. to the SQ of such AAs in a protein antigen or allergen where  
 CC the greatest local average hydrophilicity is found. Pref. the AAs in  
 CC the peptide do not exceed 50 residues, and they esp. contain 12-18  
 CC residues. The alkyl or alkenyl gp. pref. contains 12-24C and it is  
 CC gp. coupled to the terminal amino gp. of the residue opt. via a CO  
 CC benenic, oleic or mycolic acid.

XX Sequence 6 AA;

Query Match 8.9%; Score 5; DB 4; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 SSKEK 12

Db 2 sskek 6  
 |||||

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 15, 2001, 12:51:10 ; Search time 17.5 seconds  
(without alignments)  
193.997 Million cell updates/sec

Title: us-09-389-000-2\_copy\_134\_189

Perfect score: 56

Sequence: 1 KKTSLHSSKEKRLRRIRKY.....NDAASVLEATVDYVKYIREK 56

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 221992

Minimum DB seq length: 0

Maximum DB seq length: 56

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_0601.\*

1: /cgnl\_9/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /cgnl\_9/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /cgnl\_9/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /cgnl\_9/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /cgnl\_9/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /cgnl\_9/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /cgnl\_9/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /cgnl\_9/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /cgnl\_9/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /cgnl\_9/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /cgnl\_9/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /cgnl\_9/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /cgnl\_9/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /cgnl\_9/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /cgnl\_9/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /cgnl\_9/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /cgnl\_9/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /cgnl\_9/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /cgnl\_9/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /cgnl\_9/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /cgnl\_9/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	25.0	15	AA179270	PHLIX peptide use
2	6	10.7	27	AA15385	Comamonas testoste
3	5	8.9	6	AA130295	Sequence of Aas 26
4	5	8.9	6	AA150267	Peptide correspond
5	5	8.9	6	AA159948	Peptide signal seq
6	5	8.9	9	AA161019	Dynorphin-like pol
7	5	8.9	9	AA184869	H2KB 75-83 immunog
8	5	8.9	10	AA141212	Peptide fragment o
9	5	8.9	10	AA195423	HLA-B2705.75-84.
10	5	8.9	10	AA183075	HLA-B2702 CTL modu
11	5	8.9	10	AA17524	T-cell modulating

12 17 AAW07517 T-cell modulating  
13 10 AAW33785 Peptide B2705.75-8  
14 20 AAY50273 Neutrophil-activat  
15 10 AAY72484 Immunosuppressive  
16 10 AAY72488 Immunosuppressive  
17 13 AAW29421 Glucose transport  
18 14 AAW10681 Guanine thymine bi  
19 20 AAY42191 Oestrogen receptor  
20 14 AAB59409 Human Class I HLA-  
21 14 AAB59448 Human Class I HLA-  
22 15 AAR33633 HIV principal neut  
23 15 CAEV env gene TMI  
24 15 AAR89150 Oestrogen receptor  
25 15 AAY42224 Peptide fragment d  
26 14 AAB67603 C242:11 MAB kappa  
27 15 AAR50266 HLA B27 hypervaria  
28 16 AAW29423 Glucose transport  
29 16 AAY14404 Peptide CDR-I1 der  
30 16 AAY95222 Anti-platelet glyc  
31 17 AAR71440 Human MHC 1 and HL  
32 16 AAR71442 Human HLA-B27-(62-  
33 17 AAR71443 Human [Phe74]-HLA-  
34 16 AAR71425 Human MHC 1 alpha  
35 17 AAR71426 Human MHC 1 alpha  
36 17 AAR71431 Human MHC 1 alpha  
37 16 AAR71432 Human MHC 1 alpha  
38 17 AAR71433 Human MHC 1 alpha  
39 17 AAW29422 Glucose transport  
40 17 AAR32583 MHC peptide repeat  
41 18 AAW32581 MHC peptide repeat  
42 17 AAR45885 Peptide membrane b  
43 19 AAR96943 p' swap 2 mutelin o  
44 20 AAP50106 Bacillus amyloliqu  
45 15 AAR62128 Ul snRNP 70K prote

## ALIGNMENTS

RESULT 1

AA179270  
ID AAY79270 standard; Peptide; 15 AA.  
AC AAY79270;  
XX  
DT 03-JUL-2000 (first entry)  
XX  
DE PHLIX peptide used to raise antibody.  
XX  
KW PHLIX; human; testis-specific; transcription factor;  
KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
KW therapy; diagnosis; vaccine; antibody.  
XX  
OS Homo sapiens.  
XX  
PN WO200012709-A2.  
XX  
PD 09-MAR-2000.  
XX  
PF 31-AUG-1999; 99WO-US20137.  
XX  
PR 31-AUG-1998; 98US-0098610.  
PR 31-OCT-1998; 98US-0106524.  
XX  
(UROC-) UROGENESYS INC.  
PA (AFAR/) AFAR D E.  
PA (HUBE/) HUBERT R S.  
PA (RAIT/) RAITANO A B.  
XX  
PI Afar DE, Hubert RS, Raitano AB;  
XX  
DR WPI; 2000-237872/20.  
XX



CC surface antigen on a target cell. The peptides are useful for  
CC displacing antibodies bound to cell surfaces to release cells that  
CC have been positively selected by antibody-mediated binding to beads  
CC or other solid support. AAY55107 to AAY55319 represent peptides used in  
CC the exemplification of the present invention.  
XX  
SQ Sequence 6 AA;

Query Match 23.5%; Score 4; DB 20; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 SLLH 7  
| | | |  
Db 3 slh 6

Search completed: September 15, 2001, 13:00:26  
Job time: 109 sec

XX Methods for treating immunoinfective cluster virus infections -  
PT utilise antibodies or fragments characteristic of auto antibodies  
PT produced by patients with rheumatic disorders  
XX PS Disclosure: Page 8; 106pp; English.  
XX This sequence is an example of an alternating acidic/basic amino  
CC acid, hydrophilic motif possibly found in nuclear protein antigens.  
CC As well as occurring in normal human proteins, the motif is found  
CC in similar form in immunoinfective cluster viruses. The motif  
CC serves as an epitope for anti-viral antibodies and also for  
CC autoantibodies which occur in high titre in patients suffering  
CC from systemic rheumatic disorders. Sera from such patients could  
CC be used for treatment of immunoinfective cluster virus (e.g. HIV,  
CC EBV, rubella virus) infections.  
XX Sequence 6 AA;  
SQ

Query Match 23.5%; Score 4; DB 15; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 RRER 17  
| | | |  
Db 1 rrer 4

RESULT 14  
AAW21203  
ID AAW21203 standard; peptide; 6 AA.  
XX  
AC AAW21203;  
XX  
DT 29-JUL-1997 (first entry)  
XX  
DE Farnesyl synthetase derived signal oligopeptide #3.  
XX  
KW Hydrophilic; signal oligopeptide; hydrophilicity maxima; vaccine; SIV;  
KW competitive inhibitor; feedback regulator; synthesis; gastrin precursor;  
KW charge; polarity; farnesyl synthetase; plasminogen activator inhibitor 1;  
KW hydroxymethylglutaryl coenzyme A reductase; glucagon precursor; rhesus;  
KW gonadoliben precursor; plasminogen activator inhibitor 2; prorenin;  
KW Alzheimer amyloid A4; corticotropin releasing factor binding protein;  
KW apolipoprotein E; herpes virus 1 glycoprotein B; HSV1; human; OMVVS;  
KW herpes virus 2 glycoprotein B; HSV2; collagenase; apolipoprotein A;  
KW Treponema pallidum membrane protein; TMPA; islet amyloid polypeptide;  
KW fibroblast MMP1; schistosoma elastase precursor; schistosomin;  
KW hepatitis delta antigen; rev protein; HIV; VILV; angiotensinogen.  
XX  
OS Homo sapiens.  
XX  
PN W09519568-A1.  
XX  
PD 20-JUL-1995.  
XX  
PF 12-JAN-1995; 95WO-US00575.  
XX  
PR 14-JAN-1994; 94US-0182248.  
XX  
PA (RATH/) RATH M.  
XX  
PI Rath M;  
XX  
DR WPI; 1995-263953/34.  
XX  
PT Identifying signal oligopeptide(s) in protein sequence(s) - shown as  
PT regions of max. hydrophilicity, used in modulating communication  
PT between protein(s)  
XX  
PS Claim 5; Page 23; 88pp; English.  
XX

CC The sequences given in AAW21201-560 represent hydrophilic signal oligo-  
CC peptides. These signal oligopeptides are localised on the surface  
CC of the protein and are represented by the hydrophilicity maxima of  
CC the protein. These peptides are enriched in charged amino acids  
CC arranged with neutral spacer amino acids. The specific signal  
CC character of these oligopeptides is determined by a characteristic  
CC combination of conformation and charge within the signal sequence.  
CC These oligopeptides may be used as vaccines in the treatment of  
CC human disease, as competitive inhibitors to prevent or reduce the  
CC metabolic action or interaction of a selected protein by blocking  
CC its specific signal sequences, or as therapeutic agents to function  
CC as feedback regulators to reduce synthesis rate of a selected protein.  
CC These peptides may be modified by omitting one or more amino acids at  
CC the N- and/or C-terminal, by substituting one or more amino acids  
CC without consideration of charge and polarity, by substituting one or  
CC more amino acids with amino acid residues with similar charge and/or  
CC polarity, by omitting one or more amino acids or a combination of these.  
XX  
SQ Sequence 6 AA;

Query Match 23.5%; Score 4; DB 16; Length 6;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 RRER 17  
| | | |  
Db 3 rrer 6

RESULT 15  
AAW55266  
ID AAY55266 standard; peptide; 6 AA.  
XX  
AC AAY55266;  
XX  
DT 07-JAN-2000 (first entry)  
XX  
DE ATCC HB 11885 monoclonal antibody 9079 releasing peptide SEQ ID NO:160.  
XX  
KW Antibody releasing peptide; CD34; hybridoma; binding; antigen;  
KW cell surface antigen; identification; haematopoietic stem cell;  
KW tumour; cancer; immune system; therapy; displacement.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
PN US968753-A.  
XX  
PD 19-OCT-1999.  
XX  
PF 07-JUN-1995; 95US-0482228.  
XX  
PR 14-JUN-1994; 94US-0259427.  
XX  
PA (NEXE-) NEXELL THERAPEUTICS INC.  
XX  
PI Guillermo R, Helgerson SL, Deans RJ, Tseng-Law J, Kobori JA;  
PI Al-Abdaly FA;  
XX  
DR WPI; 1999-590399/50.  
XX  
PT Short peptides useful for displacing antibodies from cell surface  
PT antigens. -  
XX  
PS Example 9; Column 32; 81pp; English.  
XX  
CC The present invention describes peptides of 4-17 amino acids which  
CC displace either the anti-CD34 monoclonal antibody designated 561, the  
CC anti-CD34 mouse monoclonal antibody produced by the hybridoma ATCC  
CC HB-11646 (designated 9069), the anti-CD34 antibody produced by  
CC hybridoma ATCC HB-11885 (9079), or the anti-human breast cancer  
CC antibody produced by hybridoma ATCC HB-11884 (9187), from a cell



PI Larsen BD, Mikkelsen JD, Neve S;  
 DR WPI; 2001-159381/16.  
 XX Novel peptide agonist of Glucagon-like peptide, useful for decreasing  
 XX the level of blood glucose and for treating diseases like diabetes,  
 XX obesity and eating disorders -  
 XX  
 PS Claim 22; Page 67; 83pp; English.  
 XX  
 XX The present sequence is a peptide conjugate comprising a peptide (X)  
 CC which is an extendin at least 90 % homologous to extendin-4, a variant of  
 CC extendin comprising 1-5 deletions at positions 34-39 or a Lys at position  
 CC 40 having a lipophilic substituent, a glucagon-like peptide (GLP-1)  
 CC (7-36) or GLP-1 (7-37) having a substitution of D-Ala, Gly or alpha-amino  
 CC isobutyric acid for Ala at position 8 and/or having a lipophilic  
 CC substituent, and Z, a peptide sequence of 4-20 amino acids covalently  
 CC bound to the variant. Each amino acid in Z is selected from A, L, S, T,  
 CC Y, N, Q, D, G, K, R, H, M, Orn, and amino acid units of formula  
 CC -NH-C(R1)(R2)-C(=O)-, R1 and R2 are selected from H, Cl-6-alkyl, phenyl  
 CC and phenyl-methyl, where Cl-6-alkyl is optionally substituted with 1-3  
 CC substituents selected from halogen, hydroxy, amino, cyano, nitro,  
 CC sulfonyl, and carboxy, and phenyl and phenylmethyl are optionally  
 CC substituted with 1-3 substituents selected from Cl-6-alkyl,  
 CC C2-6-alkenyl, halogen, hydroxy, amino, cyano, nitro, sulfonyl, and  
 CC carboxy; or R1 and R2, together with the carbon atom to which they are  
 CC bound, form a cyclopentyl, cyclohexyl, or cycloheptyl ring, e.g.  
 CC 2,4-diaminobutanoic acid and 2,3-diaminopropanoic acid or its salt, or  
 CC the C-terminal amide of the peptide conjugate with the proviso that X is  
 CC not extendin-4 or extendin-3. The peptide conjugate is useful in the  
 CC manufacture of a pharmaceutical composition for use in treatment  
 CC of type 1 or type 2 diabetes, insulin resistance syndrome, obesity,  
 CC eating disorder, hyperglycaemia, metabolic disorders and gastric  
 CC disease. It is useful for treating disease states associated with  
 CC elevated blood glucose levels elicited by hormones known to increase  
 CC blood glucose levels, such as catechol amines including adrenalin,  
 CC glucocorticoids, growth hormone and glucagon. It is useful in regulation  
 CC of gastric emptying, for stimulating insulin release, for lowering plasma  
 CC lipid level, and for reducing mortality and morbidity after myocardial  
 CC infarction.  
 XX  
 SQ Sequence 40 AA;  
 Query Match 71.4%; Score 5; DB 22; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 VKGRK 7  
 DB 27 vkgrk 31  
 RESULT 15  
 AAB69975  
 ID AAB69975 standard; Peptide; 42 AA.  
 XX  
 AC AAB69975;  
 XX  
 XX 02-MAY-2001 (first entry)  
 DT  
 XX (Lys)6-Gly8-GLP-1(7-36)-(Lys)6-NH2.  
 DE  
 XX Glucagon-like peptide-1; GLP-1; extendin-4; antidiabetic; anorectic;  
 KW antinflammatory; peptide conjugate; diabetes; obesity;  
 KW insulin resistance syndrome; eating disorder; hyperglycaemia;  
 KW metabolic disorder; gastric disease; myocardial infarction.  
 XX  
 OS Synthetic.  
 XX  
 XX WO200104156-A1.  
 PN  
 XX 18-JAN-2001.  
 PD

XX 12-JUL-2000; 2000WO-DK00393.  
 XX  
 XX 12-JUL-1999; 99US-0143591.  
 PR 09-AUG-1999; 99EP-0610043.  
 XX  
 XX (ZEAL-) ZEALAND PHARM AS.  
 XX  
 XX Larsen BD, Mikkelsen JD, Neve S;  
 XX WPI; 2001-159381/16.  
 DR Novel peptide agonist of Glucagon-like peptide, useful for decreasing  
 XX the level of blood glucose and for treating diseases like diabetes,  
 XX obesity and eating disorders -  
 XX  
 PS Claim 22; Page 67; 83pp; English.  
 XX  
 XX The present sequence is a peptide conjugate comprising a peptide (X)  
 CC which is an extendin at least 90 % homologous to extendin-4, a variant of  
 CC extendin comprising 1-5 deletions at positions 34-39 or a Lys at position  
 CC 40 having a lipophilic substituent, a glucagon-like peptide (GLP-1)  
 CC (7-36) or GLP-1 (7-37) having a substitution of D-Ala, Gly or alpha-amino  
 CC isobutyric acid for Ala at position 8 and/or having a lipophilic  
 CC substituent, and Z, a peptide sequence of 4-20 amino acids covalently  
 CC bound to the variant. Each amino acid in Z is selected from A, L, S, T,  
 CC Y, N, Q, D, G, K, R, H, M, Orn, and amino acid units of formula  
 CC -NH-C(R1)(R2)-C(=O)-, R1 and R2 are selected from H, Cl-6-alkyl, phenyl  
 CC and phenyl-methyl, where Cl-6-alkyl is optionally substituted with 1-3  
 CC substituents selected from halogen, hydroxy, amino, cyano, nitro,  
 CC sulfonyl, and carboxy, and phenyl and phenylmethyl are optionally  
 CC substituted with 1-3 substituents selected from Cl-6-alkyl,  
 CC C2-6-alkenyl, halogen, hydroxy, amino, cyano, nitro, sulfonyl, and  
 CC carboxy; or R1 and R2, together with the carbon atom to which they are  
 CC bound, form a cyclopentyl, cyclohexyl, or cycloheptyl ring, e.g.  
 CC 2,4-diaminobutanoic acid and 2,3-diaminopropanoic acid or its salt, or  
 CC the C-terminal amide of the peptide conjugate with the proviso that X is  
 CC not extendin-4 or extendin-3. The peptide conjugate is useful in the  
 CC manufacture of a pharmaceutical composition for use in treatment  
 CC of type 1 or type 2 diabetes, insulin resistance syndrome, obesity,  
 CC eating disorder, hyperglycaemia, metabolic disorders and gastric  
 CC disease. It is useful for treating disease states associated with  
 CC elevated blood glucose levels elicited by hormones known to increase  
 CC blood glucose levels, such as catechol amines including adrenalin,  
 CC glucocorticoids, growth hormone and glucagon. It is useful in regulation  
 CC of gastric emptying, for stimulating insulin release, for lowering plasma  
 CC lipid level, and for reducing mortality and morbidity after myocardial  
 CC infarction.  
 XX  
 SQ Sequence 42 AA;  
 Query Match 71.4%; Score 5; DB 22; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 VKGRK 7  
 DB 33 vkgrk 37  
 Search completed: September 15, 2001, 12:47:15  
 Job time: 165 sec

CC substituent, and Z, a peptide sequence of 4-20 amino acids covalently  
 CC bound to the variant. Each amino acid in Z is selected from A, L, S, T,  
 CC Y, N, Q, D, G, K, R, H, M, Orn, and amino acid units of formula  
 CC -NH-C(R1)(R2)-C(=O)-. R1 and R2 are selected from H, Cl-6-alkyl, phenyl  
 CC and phenyl-methyl, where Cl-6-alkyl is optionally substituted with 1-3  
 CC substituents selected from halogen, hydroxy, amino, cyano, nitro,  
 CC sulfonyl, and carboxy, and phenyl and phenylmethyl are optionally  
 CC substituted with 1-3 substituents selected from Cl-6-alkyl,  
 CC C2-6-alkenyl, halogen, hydroxy, amino, cyano, nitro, sulfonyl, and  
 CC carboxy; or R1 and R2, together with the carbon atom to which they are  
 CC bound, form a cyclopentyl, cyclohexyl, or cycloheptyl ring, e.g.  
 CC 2,4-diaminobutanoic acid and 2,3-diaminopropanoic acid or its salt, or  
 CC the C-terminal amide of the peptide conjugate with the proviso that X is  
 CC not extendin-4 or extendin-3. The peptide conjugate is useful in the  
 CC manufacture of a pharmaceutical composition for use in treatment  
 CC of type 1 or type 2 diabetes, insulin resistance syndrome, obesity,  
 CC eating disorder, hyperglycaemia, metabolic disorders and gastric  
 CC disease. It is useful for treating disease states associated with  
 CC elevated blood glucose levels elicited by hormones known to increase  
 CC blood glucose levels, such as catechol amines including adrenalin,  
 CC glucocorticoids, growth hormone and glucagon. It is useful in regulation  
 CC of gastric emptying, for stimulating insulin release, for lowering plasma  
 CC lipid level, and for reducing mortality and morbidity after myocardial  
 CC infarction.

XX Sequence 38 AA;

Query Match 71.4%; Score 5; DB 22; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
 |||||  
 DB 27 vkgrk 31

RESULT 13  
 AAB69977  
 ID AAB69977 standard; Peptide; 38 AA.

XX AC AAB69977;  
 XX DT 02-MAY-2001 (first entry)  
 XX DE Gly8-GLP-1(7-36)-(Lys)8-NH2.  
 KW Glucagon-like peptide-1; GLP-1; extendin-4; antidiabetic; anorectic;  
 KW antiinflammatory; peptide conjugate; diabetes; obesity;  
 KW insulin resistance syndrome; eating disorder; hyperglycaemia;  
 KW metabolic disorder; gastric disease; myocardial infarction.

XX OS Synthetic.  
 XX PN WO200104156-A1.

XX PD 18-JAN-2001.

XX PF 12-JUL-2000; 2000WO-DK00393.

XX PR 12-JUL-1999; 99US-0143591.

XX PR 09-AUG-1999; 99EP-0610043.

XX PA (ZEAL-) ZEALAND PHARM AS.

XX PI Larsen BD, Mikkelsen JD, Neve S;

XX DR WPI; 2001-159381/16.

XX PT Novel peptide agonist of Glucagon-like peptide, useful for decreasing  
 XX the level of blood glucose and for treating diseases like diabetes,  
 XX obesity and eating disorders

XX

PS Claim 22; Page 67; 83pp; English.

XX The present sequence is a peptide conjugate comprising a peptide (X)  
 CC which is an extendin at least 90 % homologous to extendin-4, a variant of  
 CC extendin comprising 1-5 deletions at positions 34-39 or a Lys at position  
 CC 40 having a lipophilic substituent, a glucagon-like peptide (GLP-1)  
 CC (7-36) or GLP-1 (7-37) having a substitution of D-Ala, Gly or alpha-amino  
 CC isobutyric acid for Ala at position 8 and/or having a lipophilic  
 CC substituent, and Z, a peptide sequence of 4-20 amino acids covalently  
 CC bound to the variant. Each amino acid in Z is selected from A, L, S, T,  
 CC Y, N, Q, D, G, K, R, H, M, Orn, and amino acid units of formula  
 CC -NH-C(R1)(R2)-C(=O)-. R1 and R2 are selected from H, Cl-6-alkyl, phenyl  
 CC and phenyl-methyl, where Cl-6-alkyl is optionally substituted with 1-3  
 CC substituents selected from halogen, hydroxy, amino, cyano, nitro,  
 CC sulfonyl, and carboxy, and phenyl and phenylmethyl are optionally  
 CC substituted with 1-3 substituents selected from Cl-6-alkyl,  
 CC C2-6-alkenyl, halogen, hydroxy, amino, cyano, nitro, sulfonyl, and  
 CC carboxy; or R1 and R2, together with the carbon atom to which they are  
 CC bound, form a cyclopentyl, cyclohexyl, or cycloheptyl ring, e.g.  
 CC 2,4-diaminobutanoic acid and 2,3-diaminopropanoic acid or its salt, or  
 CC the C-terminal amide of the peptide conjugate with the proviso that X is  
 CC not extendin-4 or extendin-3. The peptide conjugate is useful in the  
 CC manufacture of a pharmaceutical composition for use in treatment  
 CC of type 1 or type 2 diabetes, insulin resistance syndrome, obesity,  
 CC eating disorder, hyperglycaemia, metabolic disorders and gastric  
 CC disease. It is useful for treating disease states associated with  
 CC elevated blood glucose levels elicited by hormones known to increase  
 CC blood glucose levels, such as catechol amines including adrenalin,  
 CC glucocorticoids, growth hormone and glucagon. It is useful in regulation  
 CC of gastric emptying, for stimulating insulin release, for lowering plasma  
 CC lipid level, and for reducing mortality and morbidity after myocardial  
 CC infarction.

XX SQ Sequence 38 AA;

Query Match 71.4%; Score 5; DB 22; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
 |||||  
 DB 27 vkgrk 31

RESULT 14  
 AAB69978  
 ID AAB69978 standard; Peptide; 40 AA.

XX AC AAB69978;

XX DT 02-MAY-2001 (first entry)

XX DE Gly8-GLP-1(7-36)-(Lys)10-NH2.

XX KW Glucagon-like peptide-1; GLP-1; extendin-4; antidiabetic; anorectic;  
 KW antiinflammatory; peptide conjugate; diabetes; obesity;  
 KW insulin resistance syndrome; eating disorder; hyperglycaemia;  
 KW metabolic disorder; gastric disease; myocardial infarction.

XX OS Synthetic.

XX PN WO200104156-A1.

XX PD 18-JAN-2001.

XX PF 12-JUL-2000; 2000WO-DK00393.

XX PR 12-JUL-1999; 99US-0143591.

XX PR 09-AUG-1999; 99EP-0610043.

XX PA (ZEAL-) ZEALAND PHARM AS.

XX

CC of surgery. The conjugation of a peptide of the invention to a  
 CC blood component via the reactive group provides increased stability in  
 CC the presence of peptidases. The peptides of the invention therefore  
 CC have a longer in vivo half-life as they are less susceptible to  
 CC proteolytic degradation. The present sequence represents an  
 CC insulintropic peptide of the invention.  
 XX  
 SQ Sequence 37 AA;

Query Match 71.4%; Score 5; DB 22; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
 |||||  
 Db 33 vkgrk 37

RESULT 11  
 AAB48812  
 ID AAB48812 standard; peptide; 37 AA.

XX AC AAB48812;

DT 09-MAR-2001 (first entry)

DE Insulintropic peptide, GLP-1(1-36)Lys37.

XX Insulintropic peptide; insulin production; GLP-1 derivative;  
 KW glucagon-like peptide 1; exendin derivative; reactive group;  
 KW peptidase stabilisation; blood protein; conjugation; type II diabetes;  
 KW insulin resistance; nervous system disorder; sedative; anxiolytic;  
 KW antidiabetic; neuroprotective; tranquiliser; anticonvulsant.

XX Synthetic.

OS Key Location/Qualifiers

FH Modified-site 37  
 FT /note= "Side chain is linked to a maleimidopropionic  
 FT acid (MPA) moiety, optionally via two AEEA  
 FT ([2-(2-amino)ethoxy]ethoxy acetic acid) linking  
 FT groups; C-terminal amide"

XX WO200069911-AL.

PN 23-NOV-2000.

PD 17-MAY-2000; 2000WO-US13563.

PF 17-MAY-1999; 99US-0134406.

PR 15-OCT-1999; 99US-0159783.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, L'Archeveque B, Ezrin AM, Holmes DL, Leblanc A;  
 PI St Pierre S;

XX WPI; 2001-025008/03.

DR Novel modified insulintropic peptides for treating diabetes, nervous  
 XX system disorders and for post surgery treatment, has reactive groups  
 PT which react with amino, hydroxy or thiol groups on blood components -

XX Claim 19; Page 50; 96pp; English.

PS The invention relates to modified insulintropic peptides (ITPs), or  
 CC derivatives thereof which comprise a reactive group which reacts with  
 CC amino groups, hydroxyl groups or thiol groups on blood components (e.g.,  
 CC serum albumin) to form a stable covalent bond. The insulintropic  
 CC peptides of the invention are derivatives of glucagon-like peptide 1  
 CC (GLP-1) or exendin and contain a reactive group such as a maleimido  
 CC group or a succinimidyl group. The peptides of the invention act by

CC stimulating the synthesis or expression of insulin. A composition  
 CC comprising a peptide of the invention is useful for treating diabetes,  
 CC particularly type II (maturity onset) diabetes. It is also useful as a  
 CC sedative; for the treatment of nervous system disorders including  
 CC anxiety, psychosis, seizures, panic attacks, hysteria and sleep  
 CC disorders; to induce an anxiolytic effect on the central nervous system  
 CC (CNS); to activate the CNS for the treatment of disorders such as  
 CC depression, memory loss and narcolepsy; and as a treatment for  
 CC insulin resistance, particularly that which occurs after certain types  
 CC of surgery. The conjugation of a peptide of the invention to a  
 CC blood component via the reactive group provides increased stability in  
 CC the presence of peptidases. The peptides of the invention therefore  
 CC have a longer in vivo half-life as they are less susceptible to  
 CC proteolytic degradation. The present sequence represents an  
 CC insulintropic peptide of the invention.  
 XX  
 SQ Sequence 37 AA;

Query Match 71.4%; Score 5; DB 22; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
 |||||  
 Db 33 vkgrk 37

RESULT 12

AAB69961

ID AAB69961 standard; Peptide; 38 AA.

XX AC AAB69961;

DT 02-MAY-2001 (first entry)

DE Gly8, Lys37(palmitoyl)-GLP-1(7-36)-(Lys)7-NH2.

XX Glucagon-like peptide-1; GLP-1; exendin-4; antidiabetic; anorectic;  
 KW antiinflammatory; peptide conjugate; diabetes; obesity;  
 KW insulin resistance syndrome; eating disorder; hyperglycaemia;  
 KW metabolic disorder; gastric disease; myocardial infarction.

XX Homo sapiens.

OS Synthetic.

XX WO200104156-A1.

PD 18-JAN-2001.

XX 12-JUL-2000; 2000WO-DK00393.

PF 12-JUL-1999; 99US-0143591.

PR 09-AUG-1999; 99EP-0610043.

XX (ZEAL-) ZEALAND PHARM AS.

XX Larsen BD, Mikkelsen JD, Neve S;

XX WPI; 2001-159381/16.

XX Novel peptide agonist of Glucagon-like peptide, useful for decreasing  
 PT the level of blood glucose and for treating diseases like diabetes,  
 PT obesity and eating disorders -

XX Claim 22; Page 67; 83pp; English.

PS The present sequence is a peptide conjugate comprising a peptide (X)  
 CC which is an exendin at least 90 % homologous to exendin-4, a variant of  
 CC exendin comprising 1-5 deletions at positions 34-39 or a Lys at position  
 CC 40 having a lipophilic substituent, a glucagon-like peptide (GLP-1)  
 CC (7-36) or GLP-1 (7-37) having a substitution of D-Ala, Gly or alpha-amino  
 CC isobutyric acid for Ala at position 8 and/or having a lipophilic

Best Local Similarity 100.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
| | | | |  
Db 27 vkgrk 31

RESULT 9  
AAB69974  
ID AAB69974 standard; Peptide; 36 AA.  
XX AAB69974;  
DT 02-MAY-2001 (first entry)  
XX  
DE Gly8, Lys26(palmitoyl)-GLP-1(7-36)-(Lys)6-NH2.  
XX  
KW Glucagon-like peptide-1; GLP-1; extendin-4; antidiabetic; anorectic;  
KW antiinflammatory; peptide conjugate; diabetes; obesity;  
KW insulin resistance syndrome; eating disorder; hyperglycaemia;  
KW metabolic disorder; gastric disease; myocardial infarction.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200104156-A1.  
XX  
PD 18-JAN-2001.  
XX  
PF 12-JUL-2000; 2000WO-DK00393.  
XX  
PR 12-JUL-1999; 99US-0143591.  
PR 09-AUG-1999; 99EP-0610043.  
XX  
PA (ZEAL-) ZEALAND PHARM AS.  
XX  
PI Larsen BD, Mikkelsen JD, Neve S;  
XX  
WPI; 2001-159381/16.  
XX  
Novel peptide agonist of Glucagon-like peptide, useful for decreasing  
the level of blood glucose and for treating diseases like diabetes,  
PT obesity and eating disorders -  
PT  
XX  
Claim 22; Page 67; 83pp; English.

The present sequence is a peptide conjugate comprising a peptide (X)  
which is an extendin at least 90 % homologous to extendin-4, a variant of  
extendin comprising 1-5 deletions at positions 34-39 or a Lys at position  
40 having a lipophilic substituent, a glucagon-like peptide (GLP-1)  
(7-36) or GLP-1 (7-37) having a substitution of D-Ala, Gly or alpha-amino  
isobutyric acid for Ala at position 8 and/or having a lipophilic  
substituent, and Z, a peptide sequence of 4-20 amino acids covalently  
bound to the variant. Each amino acid in Z is selected from A, L, S, T,  
Y, N, Q, D, G, K, R, H, M, Orn, and amino acid units of formula  
-NH-C(R1)(R2)-C(=O)-. R1 and R2 are selected from H, C1-6-alkyl, phenyl  
and phenyl-methyl, where C1-6-alkyl is optionally substituted with 1-3  
substituents selected from halogen, hydroxy, amino, cyano, nitro,  
sulfonyl, and carboxy, and phenyl and phenylmethyl are optionally  
substituted with 1-3 substituents selected from C1-6-alkyl,  
C2-6-alkenyl, halogen, hydroxy, amino, cyano, nitro, sulfonyl, and  
carboxy; or R1 and R2, together with the carbon atom to which they are  
bound, form a cyclopentyl, cyclohexyl, or cycloheptyl ring, e.g.  
2,4-diaminobutanoic acid and 2,3-diaminopropanoic acid or its salt, or  
the C-terminal amide of the peptide conjugate with the proviso that X is  
not extendin-4 or extendin-3. The peptide conjugate is useful in the  
manufacture of a pharmaceutical composition for use in treatment  
of type 1 or type 2 diabetes, insulin resistance syndrome, obesity,  
eating disorder, hyperglycaemia, metabolic disorders and gastric  
disease. It is useful for treating disease states associated with  
elevated blood glucose levels elicited by hormones known to increase  
blood glucose levels, such as catechol amines including adrenalin,

CC glucocorticoids, growth hormone and glucagon. It is useful in regulation  
CC of gastric emptying, for stimulating insulin release, for lowering plasma  
CC lipid level, and for reducing mortality and morbidity after myocardial  
CC infarction.  
XX  
SQ Sequence 36 AA:

Query Match 71.4%; Score 5; DB 22; Length 36;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
| | | | |  
Db 27 vkgrk 31

RESULT 10  
AAB48805  
ID AAB48805 standard; peptide; 37 AA.  
XX AAB48805;  
AC AAB48805;  
XX  
DT 09-MAR-2001 (first entry)  
XX  
DE Insulinotropic peptide, SEQ ID NO:16.  
XX  
KW Insulinotropic peptide; insulin production; GLP-1 derivative;  
KW glucagon-like peptide 1; extendin derivative; reactive group;  
KW peptidase stabilisation; blood protein; conjugation; type II diabetes;  
KW insulin resistance; nervous system disorder; sedative; anxiolytic;  
KW antidiabetic; neuroprotective; tranquiliser; anticonvulsant.  
XX  
OS Synthetic.  
XX  
PN WO200069911-A1.  
XX  
PD 23-NOV-2000.  
XX  
PF 17-MAY-2000; 2000WO-US13563.  
XX  
PR 17-MAY-1999; 99US-0134406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
PA (CONJ-) CONJUCHEM INC.  
XX  
PI Bridon DP, L'Archeveque B, Ezrin AM, Holmes DL, Leblanc A;  
PI St Pierre S;  
XX  
WPI; 2001-025008/03.  
XX  
Novel modified insulinotropic peptides for treating diabetes, nervous  
system disorders and for post surgery treatment, has reactive groups  
PT which react with amino, hydroxy or thiol groups on blood components -  
PT  
XX  
Claim 5; Page 90; 96pp; English.

The invention relates to modified insulinotropic peptides (IPs), or  
derivatives thereof which comprise a reactive group which reacts with  
amino groups, hydroxyl groups or thiol groups on blood components (e.g.,  
serum albumin) to form a stable covalent bond. The insulinotropic  
peptides of the invention are derivatives of glucagon-like peptide 1  
(GLP-1) or extendin and contain a reactive group such as a maleimido  
group or a succinimidyl group. The peptides of the invention act by  
stimulating the synthesis or expression of insulin. A composition  
comprising a peptide of the invention is useful for treating diabetes,  
particularly type II (maturity onset) diabetes. It is also useful as a  
sedative; for the treatment of nervous system disorders including  
anxiety, psychosis, seizures, panic attacks, hysteria and sleep  
disorders; to induce an anxiolytic effect on the central nervous system  
(CNS); to activate the CNS for the treatment of disorders such as  
depression, memory loss and narcolepsy; and as a treatment for  
insulin resistance, particularly that which occurs after certain types

XX Glucagon-like peptide-1; GLP-1; exendin-4; antidiabetic; anorectic;  
 KW antinflammatory; peptide conjugate; diabetes; obesity;  
 KW insulin resistance syndrome; eating disorder; hyperglycaemia;  
 KW metabolic disorder; gastric disease; myocardial infarction.  
 XX Synthetic.  
 OS  
 XX WO200104156-A1.  
 PN  
 XX 18-JAN-2001.  
 XX  
 XX 12-JUL-2000; 2000WO-DK00393.  
 XX  
 PR 12-JUL-1999; 99US-0143591.  
 PR 09-AUG-1999; 99EP-0610043.  
 XX  
 XX (ZEAL-) ZEALAND PHARM AS.  
 PA  
 XX Larsen BD, Mikkelsen JD, Neve S;  
 XX  
 XX WPI; 2001-159381/16.  
 DR  
 XX Novel peptide agonist of Glucagon-like peptide, useful for decreasing  
 PT the level of blood glucose and for treating diseases like diabetes,  
 PT obesity and eating disorders -  
 XX  
 XX Claim 22; Page 67; 83pp; English.  
 PS  
 XX The present sequence is a peptide conjugate comprising a peptide (X)  
 CC which is an exendin at least 90 % homologous to exendin-4, a variant of  
 CC exendin comprising 1-5 deletions at positions 34-39 or a Lys at position  
 CC 40 having a lipophilic substituent, a glucagon-like peptide (GLP-1)  
 CC (7-36) or GLP-1 (7-37) having a substitution of D-Ala, Gly or alpha-amino  
 CC isobutyric acid for Ala at position 8 and/or having a lipophilic  
 CC substituent, and Z, a peptide sequence of 4-20 amino acids covalently  
 CC bound to the variant. Each amino acid in Z is selected from A, L, S, T,  
 CC Y, N, Q, D, G, K, R, H, M, Orn, and amino acid units of formula  
 CC -NH-C(R1)(R2)-C(=O)-. R1 and R2 are selected from H, Cl-6-alkyl, phenyl  
 CC and phenyl-methyl, where Cl-6-alkyl is optionally substituted with 1-3  
 CC substituents selected from halogen, hydroxy, amino, cyano, nitro,  
 CC sulfonyl, and carboxy, and phenyl and phenylmethyl are optionally  
 CC substituted with 1-3 substituents selected from Cl-6-alkyl,  
 CC C2-6-alkenyl, halogen, hydroxy, amino, cyano, nitro, sulfonyl, and  
 CC carboxy; or R1 and R2, together with the carbon atom to which they are  
 CC bound, form a cyclopentyl, cyclohexyl, or cycloheptyl ring, e.g.  
 CC 2,4-diaminobutanoic acid and 2,3-diaminopropanoic acid or its salt, or  
 CC the C-terminal amide of the peptide conjugate with the proviso that X is  
 CC not exendin-4 or exendin-3. The peptide conjugate is useful in the  
 CC manufacture of a pharmaceutical composition for use in treatment  
 CC of type 1 or type 2 diabetes, insulin resistance syndrome, obesity,  
 CC eating disorder, hyperglycaemia, metabolic disorders and gastric  
 CC disease. It is useful for treating disease states associated with  
 CC elevated blood glucose levels elicited by hormones known to increase  
 CC blood glucose levels, such as catechol amines including adrenalin,  
 CC glucocorticoids, growth hormone and glucagon. It is useful in regulation  
 CC of gastric emptying, for stimulating insulin release, for lowering plasma  
 CC lipid level, and for reducing mortality and morbidity after myocardial  
 CC infarction.  
 XX Sequence 36 AA;

Query Match 71.4%; Score 5; DB 22; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
 |||||  
 DB 27 vkgrk 31

RESULT 8

AAB69973

ID AAB69973 standard; Peptide; 36 AA.  
 XX  
 AC AAB69973;  
 XX  
 XX 02-MAY-2001 (first entry)  
 XX  
 XX Gly8, Lys34(palmitoyl)-GLP-1(7-36)-(Lys)6-NH2.  
 XX  
 KW Glucagon-like peptide-1; GLP-1; exendin-4; antidiabetic; anorectic;  
 KW antinflammatory; peptide conjugate; diabetes; obesity;  
 KW insulin resistance syndrome; eating disorder; hyperglycaemia;  
 KW metabolic disorder; gastric disease; myocardial infarction.  
 XX  
 XX Homo sapiens.  
 OS Synthetic.  
 OS  
 XX WO200104156-A1.  
 PN  
 XX 18-JAN-2001.  
 XX  
 XX 12-JUL-2000; 2000WO-DK00393.  
 XX  
 PR 12-JUL-1999; 99US-0143591.  
 PR 09-AUG-1999; 99EP-0610043.  
 XX  
 XX (ZEAL-) ZEALAND PHARM AS.  
 PA  
 XX Larsen BD, Mikkelsen JD, Neve S;  
 XX  
 XX WPI; 2001-159381/16.  
 DR  
 XX Novel peptide agonist of Glucagon-like peptide, useful for decreasing  
 PT the level of blood glucose and for treating diseases like diabetes,  
 PT obesity and eating disorders -  
 XX  
 XX Claim 22; Page 67; 83pp; English.  
 PS  
 XX The present sequence is a peptide conjugate comprising a peptide (X)  
 CC which is an exendin at least 90 % homologous to exendin-4, a variant of  
 CC exendin comprising 1-5 deletions at positions 34-39 or a Lys at position  
 CC 40 having a lipophilic substituent, a glucagon-like peptide (GLP-1)  
 CC (7-36) or GLP-1 (7-37) having a substitution of D-Ala, Gly or alpha-amino  
 CC isobutyric acid for Ala at position 8 and/or having a lipophilic  
 CC substituent, and Z, a peptide sequence of 4-20 amino acids covalently  
 CC bound to the variant. Each amino acid in Z is selected from A, L, S, T,  
 CC Y, N, Q, D, G, K, R, H, M, Orn, and amino acid units of formula  
 CC -NH-C(R1)(R2)-C(=O)-. R1 and R2 are selected from H, Cl-6-alkyl, phenyl  
 CC and phenyl-methyl, where Cl-6-alkyl is optionally substituted with 1-3  
 CC substituents selected from halogen, hydroxy, amino, cyano, nitro, sulfonyl,  
 CC and carboxy, and phenyl and phenylmethyl are optionally  
 CC substituted with 1-3 substituents selected from Cl-6-alkyl,  
 CC C2-6-alkenyl, halogen, hydroxy, amino, cyano, nitro, sulfonyl, and  
 CC carboxy; or R1 and R2, together with the carbon atom to which they are  
 CC bound, form a cyclopentyl, cyclohexyl, or cycloheptyl ring, e.g.  
 CC 2,4-diaminobutanoic acid and 2,3-diaminopropanoic acid or its salt, or  
 CC the C-terminal amide of the peptide conjugate with the proviso that X is  
 CC not exendin-4 or exendin-3. The peptide conjugate is useful in the  
 CC manufacture of a pharmaceutical composition for use in treatment  
 CC of type 1 or type 2 diabetes, insulin resistance syndrome, obesity,  
 CC eating disorder, hyperglycaemia, metabolic disorders and gastric  
 CC disease. It is useful for treating disease states associated with  
 CC elevated blood glucose levels elicited by hormones known to increase  
 CC blood glucose levels, such as catechol amines including adrenalin,  
 CC glucocorticoids, growth hormone and glucagon. It is useful in regulation  
 CC of gastric emptying, for stimulating insulin release, for lowering plasma  
 CC lipid level, and for reducing mortality and morbidity after myocardial  
 CC infarction.  
 XX Sequence 36 AA;

Query Match

71.4%; Score 5; DB 22; Length 36;



XX PD 23-NOV-2000.  
 XX PF 17-MAY-2000; 2000WO-US13563.  
 XX PR 17-MAY-1999; 99US-0134406.  
 XX PR 15-OCT-1999; 99US-0159783.  
 XX PA (CONJ-) CONJUCHEM INC.  
 XX PI Bridon DP, L'Archeveque B, Ezrin AM, Holmes DL, Leblanc A;  
 XX PI St Pierre S;  
 XX DR WPI; 2001-025008/03.  
 XX Novel modified insulinotropic peptides for treating diabetes, nervous  
 PT system disorders and for post surgery treatment, has reactive groups  
 PT which react with amino, hydroxy or thiol groups on blood components -  
 XX  
 PS Claim 19; Page 57; 96pp; English.  
 XX The invention relates to modified insulinotropic peptides (ITPs), or  
 CC derivatives thereof which comprise a reactive group which reacts with  
 CC amino groups, hydroxyl groups or thiol groups on blood components (e.g.,  
 CC serum albumin) to form a stable covalent bond. The insulinotropic  
 CC peptides of the invention are derivatives of glucagon-like peptide 1  
 CC (GLP-1) or exendin and contain a reactive group such as a maleimido  
 CC group or a succinimidyl group. The peptides of the invention act by  
 CC stimulating the synthesis or expression of insulin. A composition  
 CC comprising a peptide of the invention is useful for treating diabetes,  
 CC particularly type II (maturity onset) diabetes. It is also useful as a  
 CC sedative; for the treatment of nervous system disorders including  
 CC anxiety, psychosis, seizures, panic attacks, hysteria and sleep  
 CC disorders; to induce an anxiolytic effect on the central nervous system  
 CC (CNS); to activate the CNS for the treatment of disorders such as  
 CC depression, memory loss and narcolepsy; and as a treatment for  
 CC insulin resistance, particularly that which occurs after certain types  
 CC of surgery. The conjugation of a peptide of the invention to a  
 CC blood component via the reactive group provides increased stability in  
 CC the presence of peptidases. The peptides of the invention therefore  
 CC have a longer in vivo half-life as they are less susceptible to  
 CC proteolytic degradation. The present sequence represents an  
 CC insulinotropic peptide of the invention.  
 XX Sequence 31 AA;  
 SQ  
 Query Match 71.4%; Score 5; DB 22; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3 VKGRK 7  
 Db 27 vkgrk 31  
 RESULT 6  
 ID AAB69960  
 AC AAB69960 standard; Peptide: 36 AA.  
 XX  
 XX AAB69960;  
 XX  
 DT 02-MAY-2001 (first entry)  
 XX  
 DE Gly8-GLP-1(7-36)-(Lys)6-NH2.  
 XX Glucagon-like peptide-1; GLP-1; exendin-4; antidiabetic; anorectic;  
 KW antinflammatory; peptide conjugate; diabetes; obesity;  
 KW insulin resistance syndrome; eating disorder; hyperglycaemia;  
 KW metabolic disorder; gastric disease; myocardial infarction.  
 XX  
 OS Synthetic.  
 XX

PN WO200104156-A1.  
 XX 18-JAN-2001.  
 XX PF 12-JUL-2000; 2000WO-DK00393.  
 XX PR 12-JUL-1999; 99US-0143591.  
 XX PR 09-AUG-1999; 99EP-0610043.  
 XX PA (ZEAL-) ZEALAND PHARM AS.  
 XX PI Larsen BD, Mikkelsen JD, Neve S;  
 XX PI WPI; 2001-159381/16.  
 XX Novel peptide agonist of Glucagon-like peptide, useful for decreasing  
 PT the level of blood glucose and for treating diseases like diabetes,  
 PT obesity and eating disorders -  
 XX  
 PS Claim 22; Page 67; 83pp; English.  
 XX The present sequence is a peptide conjugate comprising a peptide (X)  
 CC which is an exendin at least 90 % homologous to exendin-4, a variant of  
 CC exendin comprising 1-5 deletions at positions 34-39 or a Lys at position  
 CC 40 having a lipophilic substituent, a glucagon-like peptide (GLP-1)  
 CC (7-36) or GLP-1 (7-37) having a substitution of D-Ala, Gly or alpha-amino  
 CC isobutyric acid for Ala at position 8 and/or having a lipophilic  
 CC substituent, and Z, a peptide sequence of 4-20 amino acids covalently  
 CC bound to the variant. Each amino acid in Z is selected from A, L, S, T,  
 CC Y, N, Q, D, G, K, R, H, M, Orn, and amino acid units of formula  
 CC -NH-C(R1)(R2)-C(=O)-. R1 and R2 are selected from H, Cl-6-alkyl, phenyl  
 CC and phenyl-methyl, where Cl-6-alkyl is optionally substituted with 1-3  
 CC substituents selected from halogen, hydroxy, amino, cyano, nitro,  
 CC sulfonyl, and carboxyl, and phenyl and phenylmethyl are optionally  
 CC substituted with 1-3 substituents selected from Cl-6-alkyl,  
 CC C2-6-alkenyl, halogen, hydroxy, amino, cyano, nitro, sulfonyl, and  
 CC carboxyl; or R1 and R2, together with the carbon atom to which they are  
 CC bound, form a cyclopentyl, cyclohexyl, or cycloheptyl ring, e.g.  
 CC 2,4-diaminobutanoic acid and 2,3-diaminopropanoic acid or its salt, or  
 CC the C-terminal amide of the peptide conjugate with the proviso that X is  
 CC not exendin-4 or exendin-3. The peptide conjugate is useful in the  
 CC manufacture of a pharmaceutical composition for use in treatment  
 CC of type 1 or type 2 diabetes, insulin resistance syndrome, obesity,  
 CC eating disorder, hyperglycaemia, metabolic disorders and gastric  
 CC disease. It is useful for treating disease states associated with  
 CC elevated blood glucose levels elicited by hormones known to increase  
 CC blood glucose levels, such as catechol amines including adrenalin,  
 CC glucocorticoids, growth hormone and glucagon. It is useful in regulation  
 CC of gastric emptying, for stimulating insulin release, for lowering plasma  
 CC lipid level, and for reducing mortality and morbidity after myocardial  
 CC infarction.  
 XX Sequence 36 AA;  
 SQ  
 Query Match 71.4%; Score 5; DB 22; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3 VKGRK 7  
 Db 27 vkgrk 31  
 RESULT 7  
 ID AAB69972  
 XX AAB69972 standard; Peptide: 36 AA.  
 AC AAB69972;  
 XX  
 XX 02-MAY-2001 (first entry)  
 DT  
 XX Gly8-GLP-1(7-37)-(Lys)6-NH2.  
 DE

CC derivatives thereof which comprise a reactive group which reacts with  
 CC amino groups, hydroxyl groups or thiol groups on blood components (e.g.,  
 CC serum albumin) to form a stable covalent bond. The insulinotropic  
 CC peptides of the invention are derivatives of glucagon-like peptide 1  
 CC (GLP-1) or extendin and contain a reactive group such as a maleimido  
 CC group or a succinimidyl group. The peptides of the invention act by  
 CC stimulating the synthesis or expression of insulin. A composition  
 CC comprising a peptide of the invention is useful for treating diabetes,  
 CC particularly type II (maturity onset) diabetes. It is also useful as a  
 CC sedative; for the treatment of nervous system disorders including  
 CC anxiety, psychosis, seizures, panic attacks, hysteria and sleep  
 CC disorders; to induce an anxiolytic effect on the central nervous system  
 CC (CNS); to activate the CNS for the treatment of disorders such as  
 CC depression, memory loss and narcolepsy; and as a treatment for  
 CC insulin resistance, particularly that which occurs after certain types  
 CC of surgery. The conjugation of a peptide of the invention to a  
 CC blood component via the reactive group provides increased stability in  
 CC the presence of peptidases. The peptides of the invention therefore  
 CC have a longer in vivo half-life as they are less susceptible to  
 CC proteolytic degradation. The present sequence represents an  
 CC insulinotropic peptide of the invention.

XX Sequence 31 AA;

QY 3 VKGRK 7  
 Db 27 vkgrk 31

RESULT 4  
 AAB48813  
 ID AAB48813 standard; peptide; 31 AA.  
 AC AAB48813;  
 XX 09-MAR-2001 (first entry)  
 DT Insulinotropic peptide, GLP-1(7-36),Lys37.  
 DE Insulinotropic peptide; insulin production; GLP-1 derivative;  
 KW glucagon-like peptide 1; extendin derivative; reactive group;  
 KW peptidase stabilisation; blood protein; conjugation; type II diabetes;  
 KW insulin resistance; nervous system disorder; sedative; anxiolytic;  
 KW antidiabetic; neuroprotective; tranquiliser; anticonvulsant.  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 31  
 FT /note= "Side chain is linked to a maleimidopropionic  
 FT acid (MPA) moiety, optionally via two AEEA  
 FT (12-(2-amino)ethoxyethoxy acetic acid) linking  
 FT groups; C-terminal amide"  
 XX  
 XX WO200069911-A1.  
 PN  
 XX  
 XX 23-NOV-2000.  
 PD  
 XX  
 XX 17-MAY-2000; 2000WO-US13563.  
 PF  
 XX 17-MAY-1999; 99US-0134406.  
 PR  
 XX 15-OCT-1999; 99US-0159783.  
 PR  
 XX (CONJ-) CONJUCHEM INC.  
 PA  
 XX Bridon DP, L'Archeveque B, Ezrin AM, Holmes DL, Leblanc A;  
 PI St Pierre S;

XX WPI; 2001-025008/03.  
 DR  
 XX Novel modified insulinotropic peptides for treating diabetes, nervous  
 PT system disorders and for post surgery treatment, has reactive groups  
 PT which react with amino, hydroxy or thiol groups on blood components  
 XX  
 XX Claim 19; Page 54; 96pp; English.  
 PS  
 XX The invention relates to modified insulinotropic peptides (ITPs), or  
 CC derivatives thereof which comprise a reactive group which reacts with  
 CC amino groups, hydroxyl groups or thiol groups on blood components (e.g.,  
 CC serum albumin) to form a stable covalent bond. The insulinotropic  
 CC peptides of the invention are derivatives of glucagon-like peptide 1  
 CC (GLP-1) or extendin and contain a reactive group such as a maleimido  
 CC group or a succinimidyl group. The peptides of the invention act by  
 CC stimulating the synthesis or expression of insulin. A composition  
 CC comprising a peptide of the invention is useful for treating diabetes,  
 CC particularly type II (maturity onset) diabetes. It is also useful as a  
 CC sedative; for the treatment of nervous system disorders including  
 CC anxiety, psychosis, seizures, panic attacks, hysteria and sleep  
 CC disorders; to induce an anxiolytic effect on the central nervous system  
 CC (CNS); to activate the CNS for the treatment of disorders such as  
 CC depression, memory loss and narcolepsy; and as a treatment for  
 CC insulin resistance, particularly that which occurs after certain types  
 CC of surgery. The conjugation of a peptide of the invention to a  
 CC blood component via the reactive group provides increased stability in  
 CC the presence of peptidases. The peptides of the invention therefore  
 CC have a longer in vivo half-life as they are less susceptible to  
 CC proteolytic degradation. The present sequence represents an  
 CC insulinotropic peptide of the invention.

XX Sequence 31 AA;

QY 3 VKGRK 7  
 Db 27 vkgrk 31

RESULT 5  
 AAB48814  
 ID AAB48814 standard; peptide; 31 AA.  
 AC AAB48814;  
 XX 09-MAR-2001 (first entry)  
 DT Insulinotropic peptide, GLP-1(7-36)-DALA2, Lys37.  
 DE Insulinotropic peptide; insulin production; GLP-1 derivative;  
 KW glucagon-like peptide 1; extendin derivative; reactive group;  
 KW peptidase stabilisation; blood protein; conjugation; type II diabetes;  
 KW insulin resistance; nervous system disorder; sedative; anxiolytic;  
 KW antidiabetic; neuroprotective; tranquiliser; anticonvulsant.  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 2 /note= "D-form residue"  
 FT Modified-site 31  
 FT /note= "Side chain is linked to a maleimidopropionic  
 FT acid (MPA) moiety, optionally via two AEEA  
 FT (12-(2-amino)ethoxyethoxy acetic acid) linking  
 FT groups; C-terminal amide"  
 XX  
 XX WO200069911-A1.

PA (AFAR/) AFAR D E.  
 PA (HUBE/) HUBERT R. S.  
 PA (RAIT/) RAITANO A B.  
 XX  
 XX Afar DE, Hubert RS, Raitano AB;  
 XX  
 XX WPI: 2000-237872/20.  
 DR N-PSDB; AAZ94275.  
 XX  
 XX Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors.  
 XX  
 XX Claim 1; Fig 2A-D; 62pp; English.  
 PS  
 XX This sequence is that of human PHELIX, a novel basic Helix Loop  
 CC Helix protein thought to act as a transcription factor. PHELIX  
 CC normally exhibits a testis-specific expression pattern but is  
 CC up-regulated in prostate and other types of cancer. The invention  
 CC provides diagnostic and therapeutic methods useful in the  
 CC management of various cancers which express PHELIX, including  
 CC prostate cancer, bladder cancer, ovarian cancer and testicular  
 CC cancer, including therapies aimed at inhibition the transcription,  
 CC translation, processing or function of PHELIX. The expression  
 CC pattern of PHELIX suggests that is an ideal target for a cancer  
 CC vaccine approach to prostate cancer. PHELIX protein can also be  
 CC used to screen for agonists and antagonists of therapeutic value  
 CC and to raise antibodies.  
 XX  
 XX Sequence 405 AA;  
 SQ

Query Match 100.0%; Score 7; DB 21; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 0.62;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYVKGK 7  
 |||||  
 DB 163 pyvkgk 169

RESULT 2  
 AAR93633  
 ID AAR93633 standard; peptide; 15 AA.  
 XX  
 AC AAR93633;  
 XX  
 XX 27-APR-1996 (first entry)  
 DT  
 XX  
 XX HIV principal neutralisation epitope binding to 2F5 antibody.  
 DE  
 XX  
 XX SPNE; selected principal neutralisation epitope; vaccine; HIV;  
 KW outer membrane proteosome; Neisseria; OMPC; AIDS; 2F5 antibody.  
 XX  
 XX Synthetic.  
 OS  
 XX  
 XX GB2282379-A.  
 PN  
 XX  
 XX 05-APR-1995.  
 PD  
 XX  
 XX 23-SEP-1994; 94GB-0019255.  
 PF  
 XX  
 XX 30-SEP-1993; 93US-0129997.  
 PR  
 XX  
 XX (MERI ) MERCK & CO INC.  
 PA  
 XX  
 XX Arnold BA, Conley AJ, Kessler JA;  
 PI  
 XX  
 XX WPI: 1995-125266/17.  
 DR  
 XX  
 XX New antigenic conjugate useful as vaccine for AIDS - comprising HIV  
 PT principal neutralisation epitope covalently linked to outer membrane  
 PT proteosome of Neisseria

XX Claim 14; Page 10; 75pp; English.  
 PS  
 XX An antigenic conjugate, useful as a vaccine for AIDS, has the formula  
 CC (SPNE)n-(OMPC), where SPNE is a selected principal neutralisation  
 CC epitope of HIV, which is one of 27 specified polypeptides (including the  
 CC present sequence) or their fragments containing at least 5 amino acids  
 CC and including the DRW or DKW region; OMPC is purified outer membrane  
 CC proteosome of Neisseria (pref. N. meningitidis); and n is 1-200,  
 CC indicating the number of SPNE moieties covalently linked to the OMPC.  
 CC The conjugates may be substituted by anions, and conjugation  
 CC may be via a bigenic spacer. The SPNE polypeptides bind an HIV broadly  
 CC neutralising monoclonal antibody (2F5 antibody) specific for the  
 CC ectodomain of HIV gp41 transmembrane glycoprotein. They were originally  
 CC identified in the screening of phage epitope libraries having randomly  
 CC generated epitope polypeptides accessible to the antibody. The library  
 CC used was library ALPHA described in AAR83295. The sequences of these  
 CC polypeptides were deduced from their corresponding DNA sequence,  
 CC determined by PCR.  
 XX  
 XX Sequence 15 AA;  
 SQ

Query Match 71.4%; Score 5; DB 16; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 8.6;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
 |||||  
 DB 10 vkgrk 14

RESULT 3  
 AAB48806  
 ID AAB48806 standard; peptide; 31 AA.  
 XX  
 AC AAB48806;  
 XX  
 XX 09-MAR-2001 (first entry)  
 DT  
 XX  
 XX Insulinotropic peptide, SEQ ID NO:17.  
 DE  
 XX  
 XX Insulinotropic peptide; insulin production; GLP-1 derivative;  
 KW glucagon-like peptide 1; exendin derivative; reactive group;  
 KW peptidase stabilisation; blood protein; conjugation; type II diabetes;  
 KW insulin resistance; nervous system disorder; sedative; anxiolytic;  
 KW antidiabetic; neuroprotective; tranquilliser; anticonvulsant.  
 XX  
 OS Synthetic.  
 XX  
 XX WO2000069911-A1.  
 PN  
 XX  
 XX 23-NOV-2000.  
 PD  
 XX  
 XX 17-MAY-2000; 2000WO-US13563.  
 PF  
 XX  
 XX 17-MAY-1999; 99US-0134406.  
 PR  
 XX  
 XX 15-OCT-1999; 99US-0159783.  
 PR  
 XX  
 XX (CONJ-) CONJUCHEM INC.  
 PA  
 XX  
 XX Bridon DP, L'Archeveque B, Ezrin AM, Holmes DL, Leblanc A;  
 PI St Pierre S;  
 XX  
 XX WPI: 2001-025008/03.  
 DR  
 XX  
 XX Novel modified insulinotropic peptides for treating diabetes, nervous  
 PT system disorders and for post surgery treatment, has reactive groups  
 PT which react with amino, hydroxy or thiol groups on blood components  
 XX  
 XX Claim 5; Page 91; 96pp; English.  
 PS  
 XX The invention relates to modified insulinotropic peptides (ITPs), or  
 CC

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:47:12 ; Search time 65.58 Seconds  
(without alignments)  
6.471 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_163\_169

Perfect score: 7

Sequence: 1 PVVKGK 7

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_0601.\*

```
1: /cgnl_9/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /cgnl_9/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /cgnl_9/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /cgnl_9/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /cgnl_9/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /cgnl_9/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /cgnl_9/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /cgnl_9/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /cgnl_9/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /cgnl_9/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /cgnl_9/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /cgnl_9/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /cgnl_9/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /cgnl_9/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /cgnl_9/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /cgnl_9/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /cgnl_9/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /cgnl_9/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /cgnl_9/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /cgnl_9/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /cgnl_9/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /cgnl_9/gcgdata/geneseq/geneseq/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	100.0	405	21	AAV79269 Human testis-speci
2	5	71.4	15	16	AAR93633 HIV principal neut
3	5	71.4	31	22	AAB48806 Insulinotropic pep
4	5	71.4	31	22	AAB48813 Insulinotropic pep
5	5	71.4	31	22	AAB48814 Insulinotropic pep
6	5	71.4	35	22	AAB69960 Gly8-GLP-1(7-36)-
7	5	71.4	36	22	AAB69972 Gly8-GLP-1(7-37)-
8	5	71.4	36	22	AAB69973 Gly8, Lys34 (palmit
9	5	71.4	36	22	AAB69974 Gly8, Lys26 (palmit
10	5	71.4	37	22	AAB48805 Insulinotropic pep
11	5	71.4	37	22	AAB48812 Insulinotropic pep

12	5	71.4	38	22	AAB69961 Gly8, Lys37 (palmit
13	5	71.4	38	22	AAB69977 Gly8-GLP-1(7-36)-
14	5	71.4	40	22	AAB69978 Gly8-GLP-1(7-36)-
15	5	71.4	42	22	(Lys)6-Gly8-GLP-1-
16	5	71.4	53	20	AAV60547 Human normal bladd
17	5	71.4	62	21	AAV95586 Prostate cancer as
18	5	71.4	69	20	AAV19604 SEQ ID NO 322 from
19	5	71.4	95	19	AAV67637 MSRV-1 virus clone
20	5	71.4	95	19	AAW85049 Amino acid sequenc
21	5	71.4	143	21	AAB34111 Human secreted pro
22	5	71.4	171	21	AAB40484 Human ORFX ORF248
23	5	71.4	177	10	AAV91917 Derived sequence f
24	5	71.4	177	11	AAV07627 Human interleukin-
25	5	71.4	177	15	AAV59919 Human interleukin-
26	5	71.4	177	17	AAV92796 Human interleukin-
27	5	71.4	177	19	AAW46467 Human interleukin-
28	5	71.4	177	20	AAV41595 Human interleukin-
29	5	71.4	177	22	AAV51092 Human interleukin-
30	5	71.4	198	21	AAV11502 SEN virus protein
31	5	71.4	213	22	AAV79625 Corynebacterium gl
32	5	71.4	264	19	AAW77717 3-isopropylmalate
33	5	71.4	283	18	AAW36025 Fragment of MSRV-1
34	5	71.4	283	19	AAW71085 Multiple sclerosis
35	5	71.4	305	22	AAB87810 Mouse T2R16 amino
36	5	71.4	311	21	AAG15046 Arabidopsis thalia
37	5	71.4	313	21	AAG28697 Arabidopsis thalia
38	5	71.4	317	22	AAB87744 Human T2R14 amino
39	5	71.4	323	21	AAG49127 Arabidopsis thalia
40	5	71.4	329	21	AAG15045 Arabidopsis thalia
41	5	71.4	333	21	AAG23218 Arabidopsis thalia
42	5	71.4	333	21	AAG48263 Arabidopsis thalia
43	5	71.4	341	21	AAG49126 Arabidopsis thalia
44	5	71.4	343	17	AAW93149 C. saccharolyticum
45	5	71.4	347	21	AAG23217 Arabidopsis thalia

#### ALIGNMENTS

```
RESULT 1
AAV79269
ID AAV79269 standard; Protein; 405 AA.
XX
AC AAV79269;
XX
DT 03-JUL-2000 (first entry)
DE Human testis-specific transcription factor PHELIX.
XX
DE PHELIX; human; testis-specific; transcription factor;
KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;
KW therapy; diagnosis; vaccine.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 134..150
XX /note= "nuclear localization signal"
XX Peptide 163..169
XX /note= "nuclear localization signal"
XX Domain 140..189
XX /note= "basic Helix-Loop-Helix domain"
XX
XX WO200012709-A2.
XX
XX 09-MAR-2000.
XX
XX 31-AUG-1999; 99WO-US20137.
XX
XX 31-AUG-1998; 98US-0098610.
XX 31-OCT-1998; 98US-0106524.
XX
XX (UROC-) UROGENESYS INC.
XX
```

**THIS PAGE BLANK (USPTO)**

RESULT 15

Q9QVD3 PRELIMINARY; PRT; 8 AA.  
ID Q9QVD3; AC  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE INOSITOL 1,4,5-TRISPHOSPHATE BINDING PROTEIN.  
OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10118;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92202192; PubMed=1313009;  
RA Kanematsu T., Takeya H., Watanabe Y., Ozaki S., Yoshida M., Koga T.,  
RA Iwanaga S., Hirata M.;  
RT "Putative inositol 1,4,5-trisphosphate binding proteins in rat brain  
RT cytosol.";  
RL J. Biol. Chem. 267:6518-6525(1992).  
SQ SEQUENCE 8 AA; 1047 MW; D72415BB06C37041 CRC64;

Query Match 28.6%; Score 2; DB 11; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RK 7  
Db 7 RK 8

Search completed: September 15, 2001, 12:59:42  
Job time: 110 sec

RT a polychaete annelid, Perinereis vancaurica.";  
 RL Comp. Biochem. Physiol. C,  
 Pharmacol. Toxicol. Endocrinol. 110:297-304(1995).  
 SQ SEQUENCE 8 AA; 989 MW; 954772CAA87B0B59 CRC64;

Query Match 28.6%; Score 2; DB 5; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PY 2  
 Db 7 PY 8

RESULT 11

Q28866 PRELIMINARY; PRT; 8 AA.

ID Q28866  
 AC Q28866;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE ACTIN PROTEIN (FRAGMENT).

GN ACTIN.  
 OS Megaptera novaeangliae (Humpback whale).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;  
 OC Balaeonopteridae; Megaptera.  
 OX NCBI\_TaxID=9773;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=94285813; PubMed=7912407;  
 RA Palumbi S.R., Baker C.S.;

RT "Contrasting population structure from nuclear intron sequences and  
 mtDNA of humpback whales.";  
 RL Mol. Biol. Evol. 11:426-435(1994).  
 DR EMBL; S73467; AAD14118.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 906 MW; 69C866D1F4177408 CRC64;

Query Match 28.6%; Score 2; DB 6; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GR 6  
 Db 1 GR 2

RESULT 12

Q9TRX8 PRELIMINARY; PRT; 8 AA.

ID Q9TRX8  
 AC Q9TRX8;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DE OSTEOPOINTIN.

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]

RP SEQUENCE.

RX MEDLINE=9128766; PubMed=1676261;  
 RA Prince C.W., Dickie D., Krumdieck C.L.;

RT "Osteopontin, a substrate for transglutaminase and factor XIII  
 activity.";  
 RL Biochem. Biophys. Res. Commun. 177:1205-1210(1991).  
 SQ SEQUENCE 8 AA; 920 MW; 05DAFAF76322D767 CRC64;

Query Match 28.6%; Score 2; DB 6; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VK 4  
 Db 3 VK 4

RESULT 13

Q9GMH3 PRELIMINARY; PRT; 8 AA.

ID Q9GMH3  
 AC Q9GMH3;  
 DT 01-MAR-2001 (TReMBLrel. 16, Created)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE ACTIN (FRAGMENT).

OS Lagenorhynchus obscurus (dusky dolphin).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;  
 OC Lagenorhynchus.  
 OX NCBI\_TaxID=27611;  
 RN [1]

RP SEQUENCE FROM N.A.

RA Hare M.P., Cipriano F., Palumbi S.R.;

RT "Slow Evolution of Genetic Monophyly in Dolphins: Implications for  
 Speciation, Systematics and Conservation.";  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF140833; AAF98686.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 962 MW; 5BD1F417740862C0 CRC64;

Query Match 28.6%; Score 2; DB 6; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GR 6  
 Db 3 GR 4

RESULT 14

Q9S8Z4 PRELIMINARY; PRT; 8 AA.

ID Q9S8Z4  
 AC Q9S8Z4;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE PHOTOSYSTEM I REACTION CENTER SUBUNIT IV, PSI-E.

OS Spinacia oleracea (Spinach).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;  
 OC Caryophyllales; Chenopodiaceae; Spinacia.

OX NCBI\_TaxID=3562;  
 RN [1]

RP SEQUENCE.

RX MEDLINE=92249324; PubMed=1374333;

RA Lagoutte B., Vallon O.;

RT "Purification and membrane topology of PSI-D and PSI-E, two subunits  
 of the photosystem I reaction center.";  
 RL Eur. J. Biochem. 205:1175-1185(1992).  
 SQ SEQUENCE 8 AA; 1082 MW; 2145BB1324069044 CRC64;

Query Match 28.6%; Score 2; DB 10; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 RK 7  
 Db 4 RK 5

RC STRAIN-R545;  
RA Roy D., Coulombe M., Perron K., Roy P.H.;  
RT "Characterization of a novel 6'-N-aminoglycoside acetyltransferase  
gene aac(6')-IIC from the integron of a Chinese *Pseudomonas aeruginosa*  
RT clinical isolate."  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF162771; AAD46628.1; -.  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 930 MW; EBD5DDDD9D1A336 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YV 3  
||  
DB 7 YV 8

RESULT 7  
Q9R7T2 PRELIMINARY; PRT; 8 AA.  
ID Q9R7T2;  
AC Q9R7T2;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
DE HYPOTHETICAL PROTEIN H10004 (FRAGMENT).  
GN YQFG.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE=97061202; PubMed=8905232;  
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
RA Ikenoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,  
RA Sanpei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
RA Yano M., Horiuchi T.;  
RT "A 718-kb DNA Sequence of *Escherichia coli* K-12 Genome Corresponding  
RT to the 12.7-28.0 min Region on the Linkage Map."  
RL DNA Res. 3:137-155(1996).  
DR EMBL; D90705; BAA35310.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 8 AA; 964 MW; DF133B1DD04B476A CRC64;

Query Match 28.6%; Score 2; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PY 2  
||  
DB 2 PY 3

RESULT 8  
Q05403 PRELIMINARY; PRT; 8 AA.  
ID Q05403;  
AC Q05403;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)  
DE DNA FOR ORF'S FROM CHROMOSOME XV (FRAGMENT).  
OS Saccharomyces cerevisiae (baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-FY1679;  
RX MEDLINE=96021609; PubMed=8533473;  
RA Zumbstein E., Pearson B.M., Kaloogeropoulos A., Schweizer M.;  
RT "A 29.425 kb segment on the left arm of yeast chromosome XV contains  
RT more than twice as many unknown as known open reading frames."  
RL Yeast 11:975-986(1995).  
DR EMBL; X83121; CAA58183.1; -.  
FT NON\_TER 8  
SQ SEQUENCE 8 AA; 879 MW; 7B5322D2C441E058 CRC64;

Query Match 28.6%; Score 2; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VK 4  
||  
DB 6 VK 7

RESULT 9  
Q9UCN4 PRELIMINARY; PRT; 8 AA.  
ID Q9UCN4;  
AC Q9UCN4;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
DE CELL-SURFACE HEPARIN/HEPARANSULFATE-BINDING PROTEIN PEPTIDE 3  
DE (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92291065; PubMed=1601862;  
RA Raboudi N., Julian J., Rohde L.H., Carson D.D.;  
RT "Identification of cell-surface heparin/heparan sulfate-binding  
RT proteins of a human uterine epithelial cell line (RL95)."  
RL J. Biol. Chem. 267:11930-11939(1992).  
DR InterPro; IPR001952; -.  
DR ProDom; PD001868; -.  
SQ SEQUENCE 8 AA; 689 MW; 80B8733DD33DD87D CRC64;

Query Match 28.6%; Score 2; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KG 5  
||  
DB 6 KG 7

RESULT 10  
Q9TWH6 PRELIMINARY; PRT; 8 AA.  
ID Q9TWH6;  
AC Q9TWH6;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
DE BIOACTIVE PEPTIDE P4=PUTATIVE ESOPHAGEAL NEUROREGULATOR.  
OS Perinereis vancaurica.  
OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;  
OC Phylloclada; Nereididae; Perinereis.  
OX NCBI\_TaxID=6355;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95323338; PubMed=7599979;  
RA Takahashi T., Furukawa Y., Muneoka Y., Matsushima O., Ikeda T.,  
RA Fujita T., Minakata H., Nomoto K.;  
RT "Isolation and characterization of four novel bioactive peptides from



```

OS Human adenovirus type 7a.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=85755;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KN T96-0620, S-1058, CL 68578;
RA Crawford-Miksza L.K., Nang R.N., Schnurr D.P.;
RT "Molecular surveillance of strain variation in adenoviruses causing
RT acute respiratory disease, AV 4 and AV 7a.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065068; AAD03668.1; -.
DR EMBL; AF065066; AAD03664.1; -.
DR EMBL; AF065067; AAD03666.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 28.6%; Score 2; DB 14; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VK 4
DB 1 VK 2

RESULT 3
QYI09 ID Q9YI09 PRELIMINARY; PRT; 7 AA.
AC Q9YI09;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE PVI CORE PROTEIN (FRAGMENT).
GN PVI.
OS Human adenovirus type 4.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z-G 95-873, RI-67, 55142;
RA Crawford-Miksza L.K., Nang R.N., Schnurr D.P.;
RT "Molecular surveillance of strain variation in adenoviruses causing
RT acute respiratory disease, AV 4 and AV 7a.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065064; AAD03659.1; -.
DR EMBL; AF065062; AAD03653.1; -.
DR EMBL; AF065063; AAD03656.1; -.
FT NON_TER 1
SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 28.6%; Score 2; DB 14; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VK 4
DB 1 VK 2

RESULT 4
Q56140 ID Q56140 PRELIMINARY; PRT; 8 AA.
AC Q56140;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE (ST11) STP6 CONTROL REGION (FRAGMENT).
GN STP6.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.

```

```

OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ST11;
RX MEDLINE=95047254; PubMed=7958782;
RA Constable A., Mollet B.;
RT "Isolation and characterisation of promoter regions from Streptococcus
RT thermophilus.";
RL FEMS Microbiol. Lett. 122:85-90(1994).
DR EMBL; X78210; CAA55045.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 846 MW; ED086772D5B045B6 CRC64;

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GR 6
DB 7 GR 8

RESULT 5
Q56429 ID Q56429 PRELIMINARY; PRT; 8 AA.
AC Q56429;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE GAPDH (FRAGMENT).
OS Thermus aquaticus (subsp. thermophilus).
OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB-8;
RX MEDLINE=89025722; PubMed=3052437;
RA Bowen D., Littlechild J.A., Fothergill J.E., Watson H.C., Hall L.;
RT "Nucleotide sequence of the phosphoglycerate kinase gene from the
RT extreme thermophile Thermus thermophilus. Comparison of the deduced
RT amino acid sequence with that of the mesophilic yeast phosphoglycerate
RT kinase.";
RL Biochem. J. 254:509-517(1988).
DR EMBL; X12464; CAA31005.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 885 MW; 33C873333732C72B CRC64;

Query Match 28.6%; Score 2; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KG 5
DB 6 KG 7

RESULT 6
Q9S443 ID Q9S443 PRELIMINARY; PRT; 8 AA.
AC Q9S443;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE BETA-LACTAMASE (FRAGMENT).
GN PSE2.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:57:52 ; Search time 17.47 seconds  
(without alignments)  
53.013 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_163\_169  
Perfect score: 7  
Sequence: 1 PYVKGRK 7

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 890

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_16:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2	28.6	7	14 Q9YVE3	Q9YVE3 human adeno
2	2	28.6	7	14 Q9YIR0	Q9YIR0 human adeno
3	2	28.6	7	14 Q9YIQ9	Q9YIQ9 human adeno
4	2	28.6	8	2 Q56140	Q56140 streptococc
5	2	28.6	8	2 Q56429	Q56429 thermus aqu
6	2	28.6	8	2 Q9S443	Q9S443 pseudomonas
7	2	28.6	8	2 Q9R7T2	Q9R7T2 escherichia
8	2	28.6	8	3 Q05403	Q05403 saccharomyc
9	2	28.6	8	4 Q9UCN4	Q9UCN4 homo sapien
10	2	28.6	8	5 Q9TWH6	Q9TWH6 perinereis
11	2	28.6	8	6 Q28866	Q28866 megaptera n
12	2	28.6	8	6 Q9TRX8	Q9TRX8 bos taurus
13	2	28.6	8	6 Q9GMH3	Q9GMH3 lagenorhync
14	2	28.6	8	10 Q9S824	Q9S824 spinacia ol
15	2	28.6	8	11 Q9QVD3	Q9QVD3 rattus sp.
16	2	28.6	8	14 Q84271	Q84271 human papil
17	2	28.6	8	14 Q84273	Q84273 human papil
18	2	28.6	8	14 Q9E807	Q9E807 beet soil-b
19	2	28.6	8	14 Q9E8Q6	Q9E8Q6 beet soil-b

20	2	28.6	8	14 Q9E8Q5	Q9E8Q5 beet soil-b
21	2	28.6	8	14 Q9E8Q4	Q9E8Q4 beet soil-b
22	2	28.6	8	14 Q9E8Q3	Q9E8Q3 beet soil-b
23	2	28.6	8	14 Q9E8Q2	Q9E8Q2 beet soil-b
24	2	28.6	8	14 Q9E8Q1	Q9E8Q1 beet soil-b
25	2	28.6	8	14 Q9E8Q0	Q9E8Q0 beet soil-b
26	2	28.6	8	14 Q9E8P9	Q9E8P9 beet soil-b
27	2	28.6	8	14 Q9E8P8	Q9E8P8 beet soil-b
28	2	28.6	8	14 Q9E8P7	Q9E8P7 beet soil-b
29	2	28.6	8	14 Q9DSN6	Q9DSN6 beet soil-b
30	2	28.6	8	14 Q9DSN5	Q9DSN5 beet soil-b
31	2	28.6	8	14 Q9DSN4	Q9DSN4 beet soil-b
32	2	28.6	8	14 Q9DSN3	Q9DSN3 beet soil-b
33	2	28.6	8	14 Q9DSN2	Q9DSN2 beet soil-b
34	2	28.6	8	14 Q9DSN1	Q9DSN1 beet soil-b
35	2	28.6	8	14 Q9DSN0	Q9DSN0 beet soil-b
36	2	28.6	9	2 Q45852	Q45852 clostridium
37	2	28.6	9	2 Q46179	Q46179 clostridium
38	2	28.6	9	2 Q9R7T6	Q9R7T6 escherichia
39	2	28.6	9	4 Q15999	Q15999 homo sapien
40	2	28.6	9	4 Q16220	Q16220 homo sapien
41	2	28.6	9	4 Q9S574	Q9S574 homo sapien
42	2	28.6	9	4 Q9UQW0	Q9UQW0 homo sapien
43	2	28.6	9	4 Q16276	Q16276 homo sapien
44	2	28.6	9	6 Q28093	Q28093 bos taurus
45	2	28.6	9	6 Q9TUY0	Q9TUY0 monodelphis

#### ALIGNMENTS

RESULT 1  
Q9YVE3  
ID Q9YVE3 PRELIMINARY; PRT; 7 AA.  
AC Q9YVE3;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE PVI CORE PROTEIN (FRAGMENT).  
GN PVI.  
OS Human adenovirus type 7.  
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
OX NCBI\_TaxID=10519;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GMEN;  
RA Crawford-Miksza L.K., Nang R.N., Schnurr D.P.;  
RT "Molecular surveillance of strain variation in adenoviruses causing acute respiratory disease, AV 4 and AV 7a."  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF065065; AAD03662.1; -  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 28.6%; Score 2; DB 14; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VK 4

DB 1 VK 2

RESULT 2

Q9YIR0 PRELIMINARY; PRT; 7 AA.

ID Q9YIR0  
AC Q9YIR0  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
DE PVI CORE PROTEIN (FRAGMENT).  
GN PVI.

---

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidea; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT CHAIN 1 8 CARCINUSTATIN 7.  
 FT CHAIN 2 8 CARCINUSTATIN 6.  
 FT CHAIN 4 8 CARCINUSTATIN 1.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 825 MW; 922879CDCB4775BD CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PY 2  
 ||  
 Db 3 PY 4

## RESULT 14

ALL8\_CARMA STANDARD; PRT; 8 AA.  
 AC P81811;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE CARCINUSTATIN 8.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidea; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 795 MW; 922879CDCB47687D CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PY 2  
 ||  
 Db 3 PY 4

## RESULT 15

ALL9\_CARMA STANDARD; PRT; 8 AA.  
 ID ALL9\_CARMA

AC P81812;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 9.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidea; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 781 MW; 7C2879CDCB476878 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PY 2  
 ||  
 Db 3 PY 4

Search completed: September 15, 2001, 12:59:59  
 Job time: 107 sec

DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CALLATOSTATIN 4 (LEU-CALLATOSTATIN 4).  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
OX NCBI\_TaxID=27454;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Thoracic ganglion;  
RX MEDLINE=93211980; PubMed=8460157;  
RA Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S.,  
RA Thorpe A.;  
RT "Callatostatins: neuropeptides from the blowfly Calliphora vomitoria  
with sequence homology to cockroach allatostatins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).  
RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=94291167; PubMed=8020069;  
RA Duve H., Thorpe A.;  
RT "Distribution and functional significance of Leu-callatostatins in  
the blowfly Calliphora vomitoria.";  
RL Cell Tissue Res. 276:367-379(1994).  
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY  
INVOLVED IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE  
BEHAVE AS POTENT INHIBITORS OF VISCERAL MUSCLES DUE TO ITS ABILITY TO  
FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.  
CC -!- TISSUE SPECIFICITY: BRAIN, SUBESOPHAGEAL GANGLION, RETROCEREAL  
COMPLEX, THORACICO-ABDOMINAL GANGLION, PERIPHERAL NEUROSECRETORY  
SYSTEM AND INTESTINE.  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
DR PIR; D47393; D47393.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 8  
FT UNSURE 1 OR N  
SQ SEQUENCE 8 AA; 954 MW; D32879D5AB47740A CRC64;  
  
Query Match 28.6%; Score 2; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 PY 2  
DB 3 PY 4  
  
RESULT 11  
ALL4\_CYPDPO STANDARD; PRT; 8 AA.  
AC P82155;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE CYDIASATIN 4.  
OS Cydia pomonella (Codling moth).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Tortricodea; Tortricidae; Olethreutinae; Cydia.  
OX NCBI\_TaxID=82600;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Larva;  
RX MEDLINE=98054539; PubMed=9392829;  
RA Duve H., Johnsen A.H., Maestros J.-L., Scott A.G., Winstanley D.,  
RA Davey M., East P.D., Thorpe A.;  
RT "Lepidopteran peptides of the allatostatin superfamily.";  
RL Peptides 18:1301-1309(1997).  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 8  
FT UNSURE 8  
SQ SEQUENCE 8 AA; 883 MW; 7D9879CAB8477768 CRC64;

SQ SEQUENCE 8 AA; 910 MW; 922879D5AB47740D CRC64;  
  
Query Match 28.6%; Score 2; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 PY 2  
DB 3 PY 4  
  
RESULT 12  
ALL5\_CALVO STANDARD; PRT; 8 AA.  
AC P41841;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CALLATOSTATIN 5 (MET-CALLATOSTATIN 1) ((HYP3)MET-CALLATOSTATIN).  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
OX NCBI\_TaxID=27454;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93211980; PubMed=8460157;  
RA Duve H., Johnsen A.H., Scott A.G., Yu C.G., Yagi K.J., Tobe S.S.,  
RA Thorpe A.;  
RT "Callatostatins: neuropeptides from the blowfly Calliphora vomitoria  
with sequence homology to cockroach allatostatins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).  
RN [2]  
RP CHARACTERIZATION, AND HYDROXYLATION.  
RC TISSUE=Head;  
RX MEDLINE=94342269; PubMed=8063725;  
RA Duve H., Johnsen A.H., Scott A.G., East P., Thorpe A.;  
RT "[Hyp3]Met-callatostatin. Identification and biological properties of  
a novel neuropeptide from the blowfly Calliphora vomitoria.";  
RL J. Biol. Chem. 269:21059-21066(1994).  
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY  
A ROLE IN THE INTEGRATION OF INFORMATION WITHIN THE BRAIN. MAY BE  
INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO  
BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO  
FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.  
CC -!- TISSUE SPECIFICITY: NEURONS WITHIN BRAIN AND ABDOMINAL GANGLION.  
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
DR PIR; E47393; E47393.  
KW Neuropeptide; Amidation; Hydroxylation.  
FT MOD\_RES 3  
FT MOD\_RES 8  
FT MOD\_RES 8  
SQ SEQUENCE 8 AA; 883 MW; 7D9879CAB8477768 CRC64;  
  
Query Match 28.6%; Score 2; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 PY 2  
DB 3 PY 4  
  
RESULT 13  
ALL7\_CARMA STANDARD; PRT; 8 AA.  
AC P81809; P81810; P81804;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CARCINUSTATIN 7 [CONTAINS: CARCINUSTATIN 6; CARCINUSTATIN 1].  
OS Carcinus maenas (Common shore crab) (Green crab).

RA Baghdassarian D.;  
 RT "A transaldolase. An enzyme implicated in crab steroidogenesis.";  
 RL Endocrine 5:23-32(1996).  
 CC -|- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED  
 CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED  
 CC IN ALL EUKARYOTIC CELLS.  
 CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -|- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:  
 CC 6.8, ITS MW IS: 46 KDA.  
 CC -|- SIMILARITY: BELONGS TO THE ACTIN FAMILY.  
 DR InterPro: IPR000279;  
 DR PROSITE: PS00406; ACTINS.1; PARTIAL.  
 DR PROSITE: PS00432; ACTINS.2; PARTIAL.  
 DR PROSITE: PS01132; ACTINS\_ACT-LIKE; PARTIAL.  
 KW Structural protein.  
 FT NON\_TER 1  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 976 MW; 1424005AB2CAAE3 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RK 7  
 ||  
 Db 7 RK 8

RESULT 7  
 ID ALL15\_CARMA STANDARD; PRT; 8 AA.  
 AC P8181;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE CARCINUSTATIN 15.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -|- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -|- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT MOD\_RES 8  
 FT MOD\_RES 8  
 SQ SEQUENCE 8 AA; 811 MW; 922879D5AB47687D CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PY 2  
 ||  
 Db 3 PY 4

RESULT 8  
 ID ALL16\_CARMA STANDARD; PRT; 8 AA.  
 AC P81819;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 16.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -|- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -|- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation; Multigene family.  
 FT MOD\_RES 8  
 FT MOD\_RES 8  
 SQ SEQUENCE 8 AA; 813 MW; 7C286B45AB476878 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PY 2  
 ||  
 Db 3 PY 4

RESULT 9  
 ID ALL3\_CYPDPO STANDARD; PRT; 8 AA.  
 AC P82154;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CYDIASTATIN 3.  
 OS Cydia pomonella (Codling moth).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Tortricoidae; Tortricidae; Olethreutinae; Cydia.  
 OX NCBI\_TaxID=82600;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Larva;  
 RX MEDLINE=98054539; PubMed=9392829;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,  
 RA Davey M., East P.D., Thorpe A.;  
 RT "Lepidopteran peptides of the allatostatatin superfamily.";  
 RL Peptides 18:1301-1309(1997).  
 CC -|- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8  
 FT MOD\_RES 8  
 SQ SEQUENCE 8 AA; 926 MW; C82879D5AB477415 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PY 2  
 ||  
 Db 3 PY 4

RESULT 10  
 ID ALL4\_CALVO STANDARD; PRT; 8 AA.  
 AC P41840;  
 DT 01-NOV-1995 (Rel. 32, Created)

RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Multigene family.  
 SQ SEQUENCE 7 AA; 796 MW; 672879CDCB476B70 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PY 2  
 ||  
 Db 2 PY 3

## RESULT 3

ALL4\_CARMA  
 ID ALL4\_CARMA STANDARD; PRT; 7 AA.  
 AC P81807;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 4.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachiura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the  
 RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Multigene family.  
 SQ SEQUENCE 7 AA; 782 MW; 672879CDCB476AC0 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PY 2  
 ||  
 Db 2 PY 3

## RESULT 4

ALL5\_CARMA  
 ID ALL5\_CARMA STANDARD; PRT; 7 AA.  
 AC P81808;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CARCINUSTATIN 5.  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachiura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
 RX MEDLINE=98121193; PubMed=9461295;  
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
 RA Thorpe A.;  
 RT "Isolation and identification of multiple neuropeptides of the

RT allatostatin superfamily in the shore crab Carcinus maenas.";  
 RL Eur. J. Biochem. 250:727-734(1997).  
 CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.  
 CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.  
 KW Neuropeptide; Multigene family.  
 FT MOD.RES 7 AMIDATION.  
 SQ SEQUENCE 7 AA; 781 MW; 672879CDCB476420 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PY 2  
 ||  
 Db 2 PY 3

## RESULT 5

GFRP\_MOUSE  
 ID GFRP\_MOUSE STANDARD; PRT; 7 AA.  
 AC P99025;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE GTP CYCLOHYDROLASE I FEEDBACK REGULATORY PROTEIN (P35) (FRAGMENT).  
 GN GCHFR OR GFRP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Liver;  
 RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,  
 RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,  
 RA Cowthorne M.;  
 RL Submitted (AUG-1998) to the SWISS-PROT data bank.  
 CC -!- FUNCTION: MEDIATES TETRAHYDROBIPTERIN INHIBITION OF GTP  
 CC CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLALANINE  
 CC (BY SIMILARITY).  
 CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 DR SWISS-2DPAGE; P99025; MOUSE.  
 FT INIT\_MET 0  
 FT NON\_TER 7  
 SQ SEQUENCE 7 AA; 806 MW; 71B5B057273B4700 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PY 2  
 ||  
 Db 1 PY 2

## RESULT 6

ACT\_CARMA  
 ID ACT\_CARMA STANDARD; PRT; 8 AA.  
 AC P80709;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE ACTIN (FRAGMENT).  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachiura; Portunoidae; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RA Lachaise F., Somme G., Carpentier G., Granjeon E., Webster S.,

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:58:12 : Search time 8.46 Seconds

(without alignments)  
28.344 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_163\_169

Perfect score: 7

Sequence: 1 PYVKGRK 7

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 331

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2	28.6	5	1 UF01_MOUSE	P38639 mus musculus
2	2	28.6	7	1 ALL3_CARMA	P81806 carcinus ma
3	2	28.6	7	1 ALL4_CARMA	P81807 carcinus ma
4	2	28.6	7	1 ALL5_CARMA	P81808 carcinus ma
5	2	28.6	7	1 GFRP_MOUSE	P99025 mus musculus
6	2	28.6	8	1 ACT_CARMA	P80709 carcinus ma
7	2	28.6	8	1 ALL5_CARMA	P81818 carcinus ma
8	2	28.6	8	1 ALL6_CARMA	P81819 carcinus ma
9	2	28.6	8	1 ALL3_CVDPO	P82154 cydia pomon
10	2	28.6	8	1 ALL4_CALVO	P41840 calliphora
11	2	28.6	8	1 ALL4_CVDPO	P82155 cydia pomon
12	2	28.6	8	1 ALL5_CALVO	P41841 calliphora
13	2	28.6	8	1 ALL7_CARMA	P81809 carcinus ma
14	2	28.6	8	1 ALL8_CARMA	P81811 carcinus ma
15	2	28.6	8	1 ALL9_CARMA	P81812 carcinus ma
16	2	28.6	8	1 B44K_PORGI	P81886 porphyron
17	2	28.6	8	1 RSL_ERWCH	P37985 erwinia chr
18	2	28.6	8	1 RST_MYCIT	P33564 mycobacteri
19	2	28.6	9	1 ALL10_CARMA	P81813 carcinus ma
20	2	28.6	9	1 CONO_CONGE	P05486 conus geogr
21	2	28.6	9	1 COXE_THUOB	P80975 thunnus obe
22	2	28.6	9	1 FIBB_ERYPA	P19346 erythrocebu
23	2	28.6	9	1 FIBB_WACFU	P19345 macaca fusc
24	2	28.6	9	1 FIBB_PAPAN	P19344 papio anubi
25	2	28.6	9	1 FIBB_PAPHA	P19343 papio hamad
26	2	28.6	9	1 FIBB_THEGE	P19342 theropitheci
27	2	28.6	9	1 FLA2_TREHY	P80159 treponema h
28	2	28.6	9	1 NEUX_HUMAN	P04277 homo sapien
29	2	28.6	9	1 NEUX_RAT	P11382 rattus norv
30	2	28.6	9	1 PGLR_DIAAB	P81179 diaprepes a
31	2	28.6	9	1 ULAD_HUMAN	P31929 homo sapien
32	2	28.6	9	1 ULAK_MOUSE	P99031 mus musculus
33	2	28.6	9	1 UPA7_HUMAN	P30093 homo sapien

34	2	28.6	10	1 ANGL_BOTJA	Q10581 bothrops ja
35	2	28.6	10	1 ANGT_BOVIN	P01017 bos taurus
36	2	28.6	10	1 ANGT_CHICK	P01018 gallus gall
37	2	28.6	10	1 APE_CAPGI	P80474 capnocytoph
38	2	28.6	10	1 RLA2_MOUSE	P99027 mus musculu
39	2	28.6	10	1 TPIS_NICPL	P19118 nicotiana p
40	2	28.6	10	1 UHA3_HUMAN	P40930 homo sapien
41	2	28.6	10	1 UPA5_HUMAN	P30091 homo sapien
42	1	14.3	3	1 GRWM_HUMAN	P01157 homo sapien
43	1	14.3	3	1 LUXE_VIBFI	P24272 vibrio fisc
44	1	14.3	3	1 THYL_PIG	P01151 sus scrofa
45	1	14.3	4	1 ACH1_ACHFU	P35904 achatina fu

ALIGNMENTS

RESULT 1  
ID UF01\_MOUSE STANDARD; PRT; 5 AA.  
AC P38639;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF FIBROBLASTS (P19) (FRAGMENT).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Fibroblast;  
RX MEDLINE=95009907; PubMed=7523108;  
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;  
RT "Separation and sequencing of familiar and novel murine proteins  
RT using preparative two-dimensional gel electrophoresis.";  
RL Electrophoresis 15:735-745(1994).  
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
FT NON\_TER 5  
SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;

Query Match 28.6%; Score 2; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 GR 6  
Db 3 GR 4

RESULT 2  
ID ALL3\_CARMA STANDARD; PRT; 7 AA.  
AC P81806;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CARCINUSTATIN 3.  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.  
OX NCBI\_TaxID=6759;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;  
RX MEDLINE=98121193; PubMed=9461295;  
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,  
RA Thorpe A.;  
RT "Isolation and identification of multiple neuropeptides of the  
RT allatostatin superfamily in the shore crab Carcinus maenas.";



Query Match 28.6%; Score 2; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RK 7  
 ||  
 Db 1 RK 2

RESULT 14  
 PT0511  
 T-cell receptor beta chain V-D-J region (100-4A) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PT0511  
 R;Feeney, A.J.  
 J. Exp. Med. 174, 115-124, 1991  
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A:Reference number: PT0509; MUID:91277601  
 A:Accession: PT0511  
 A>Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-6 <FEE>  
 A:Experimental source: adult thymus, strain BALB/c  
 C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KG 5  
 ||  
 Db 5 KG 6

RESULT 15  
 PT0516  
 T-cell receptor beta chain V-D-J region (100-4AF) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
 C:Accession: PT0516  
 R;Feeney, A.J.  
 J. Exp. Med. 174, 115-124, 1991  
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
 A:Reference number: PT0509; MUID:91277601  
 A:Accession: PT0516  
 A>Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-6 <FEE>  
 A:Experimental source: adult thymus, strain BALB/c  
 C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KG 5  
 ||  
 Db 5 KG 6

Search completed: September 15, 2001, 12:59:15  
 Job time: 104 sec

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0553

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601

A:Accession: PT0553

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-5 <PEE>

A:Experimental source: day 18 fetal thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.2e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GR 6

Db 3 GR 4

RESULT 9

A35890

RNA-directed DNA polymerase (EC 2.7.7.49) 66K chain - human immunodeficiency virus type

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 31-Dec-1993

C:Accession: A35890

R:Batthurst, I.C.; Moen, L.K.; Lujan, M.A.; Gibson, H.L.; Feucht, P.H.; Pichuanes, S.; O

Biochem. Biophys. Res. Commun. 171, 589-595, 1990

A:Title: Characterization of the human immunodeficiency virus type-1 reverse transcripta

A:Reference number: A35890; MUID:90386627

A:Accession: A35890

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-6 <BAT>

C:Keywords: nucleotidyltransferase

Query Match 28.6%; Score 2; DB 2; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.2e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RK 7

Db 3 RK 4

RESULT 10

S11024

Hydrogensulfite reductase (EC 1.8.99.3) chain 1 - Desulfovibrio thermophilus (fragment)

N:Alternate names: bisulfite reductase; desulfosulfidin

C:Species: Desulfovibrio thermophilus

C:Date: 19-Mar-1997 #sequence\_revision 30-Jan-1998 #text\_change 30-Jan-1998

C:Accession: S11024

R:Faure, G.; Lino, A.R.; Czechowski, M.; Kang, L.; DerVartanian, D.V.; Moura, J.J.G.; L

Biochim. Biophys. Acta 1040, 112-118, 1990

A:Title: Purification and characterization of bisulfite reductase (desulfosulfidin) from

A:Reference number: S11024; MUID:90335276

A:Accession: S11024

A:Molecule type: protein

A:Residues: 1-6 <FAU>

C:Keywords: oxidoreductase

Query Match 28.6%; Score 2; DB 2; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.2e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VK 4

Db 3 VK 4

RESULT 11

A60494

antineoplastic glycoprotein - sea hare (Dolabella auricularia) (fragment)

N:Alternate names: dolabellarin C

C:Species: Dolabella auricularia

C:Date: 19-Mar-1993 #sequence\_revision 19-Mar-1993 #text\_change 18-Jun-1993

C:Accession: A60494

R:Kisugi, J.; Kamiya, H.; Yamazaki, M.

Dev. Comp. Immunol. 13, 3-8, 1989

A:Title: Purification of dolabellarin-C an antineoplastic glycoprotein in the body fl

A:Reference number: A60494; MUID:89357188

A:Accession: A60494

A:Molecule type: protein

A:Residues: 1-6 <KIS>

C:Keywords: cytolysis; glycoprotein; trimer

Query Match 28.6%; Score 2; DB 2; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.2e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GR 6

Db 4 GR 5

RESULT 12

I51434

H4 histone - African clawed frog (fragment)

C:Species: Xenopus laevis (African clawed frog)

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C:Accession: I51434

R:Woodland, H.R.; Warmington, J.R.; Ballantine, J.E.M.; Turner, P.C.

Nucleic Acids Res. 12, 4939-4958, 1984

A:Title: Are there major developmentally regulated H4 gene classes in Xenopus?

A:Reference number: I51391; MUID:84247348

A:Accession: I51434

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-6 <WOO>

A:Cross-references: GB:K02304; NID:g214227; PIDN:AAA49738.1; PID:g555517

Query Match 28.6%; Score 2; DB 2; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.2e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GR 6

Db 3 GR 4

RESULT 13

I48126

alpha-tubulin - Chinese hamster (fragment)

C:Species: Cricetulus griseus (Chinese hamster)

C:Date: 04-Sep-1997 #sequence\_revision 07-Nov-1997 #text\_change 05-Nov-1999

C:Accession: I48126

R:Elliot, E.M.; Sarangi, F.; Henderson, G.; Ling, V.

Can. J. Biochem. Cell Biol. 63, 511-518, 1985

A:Title: Cloning of 11 alpha-tubulin gene sequences from the genome of Chinese hamste

A:Reference number: I48126; MUID:86001952

A:Accession: I48126

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-6 <RES>

A:Cross-references: GB:M25895; NID:g341417; PIDN:AAA74493.1; PID:g516601

C:Genetics:

A:Introns: 3/3

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RK 7  
||  
Db 3 RK 4

RESULT 3  
A:Accession: A25844  
A:Title: auto-RF amide neuropeptide - sea pansy (Renilla koellikeri)  
C:Species: Renilla koellikeri (Koelliker's sea pansy)  
C:Date: 21-May-1988 #sequence\_revision 30-Sep-1993 #text\_change 11-Jul-1997  
C:Accession: A25844  
R:Grimmelikhuijzen, C.J.P.; Groeger, A.  
FEBS Lett. 211, 103-108, 1987  
A:Title: Isolation of the neuropeptide pGlu-Gly-Arg-Phe-amide from the pennatulid Renilla koellikeri  
A:Reference number: A25844  
A:Accession: A25844  
A:Molecule type: protein  
A:Residues: 1-4 <GRI>  
C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:4/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 28.6%; Score 2; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GR 6  
||  
Db 2 GR 3

RESULT 4  
I40702  
A:Title: Citrobacter diversus (fragment)  
C:Species: Citrobacter diversus  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 16-Aug-1996  
C:Accession: I40702  
R:Versalovic, J.; Lupski, J.R.  
Mol. Microbiol. 8, 343-355, 1993  
A:Title: Conservation and evolution of the rpsu-dnaG-rpod macromolecular synthesis (MMS)  
A:Reference number: I40702; MUID:93302510  
A:Accession: I40702  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-5 <RES>  
C:Cross-references: GB:L01754; NID:g144439  
C:Genetics:  
A:Gene: dnaG

Query Match 28.6%; Score 2; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GR 6  
||  
Db 3 GR 4

RESULT 5  
PQ0009  
A:Title: angiotensin-converting enzyme inhibitor (PLP-2) - common fig  
N:Alternate names: ficus latex peptide 2  
C:Species: Ficus carica (Common fig)  
C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 08-Dec-1995  
C:Accession: PQ0009  
R:Maruyama, S.; Miyoshi, S.; Tanaka, H.  
Agric. Biol. Chem. 53, 2763-2767, 1989  
A:Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.  
A:Reference number: PQ0008

Query Match 28.6%; Score 2; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GR 6  
||  
Db 3 GR 4

RESULT 8  
PT0553  
A:Title: T-cell receptor beta chain V-D-J region (126-1C) - mouse (fragment)  
C:Species: Mus musculus (house mouse)

Query Match 28.6%; Score 2; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GR 6  
||  
Db 3 GR 4

RESULT 7  
PT0608  
A:Title: T-cell receptor beta chain V-D-J region (120-2CF) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0608  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions  
A:Reference number: PT0509; MUID:91277601  
A:Accession: PT0608  
A>Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-5 <FEE>  
A:Experimental source: newborn thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 28.6%; Score 2; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GR 6  
||  
Db 3 GR 4

RESULT 6  
SS3595  
A:Title: hypothetical protein (upstream of transcription factor, CCAAT-binding) - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 07-May-1999  
C:Accession: SS3595  
R:Calhoun, C.F.; Bouwman, P.R.J.; Snippe, L.; Ab, G.  
Nucleic Acids Res. 22, 5540-5547, 1994  
A:Title: Translation start site multiplicity of the CCAAT/enhancer binding protein alpha  
A:Reference number: SS3595; MUID:95140613  
A:Accession: SS3595  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-5 <CAL>  
A:Cross-references: EMBL:X66844

Query Match 28.6%; Score 2; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GR 6  
||  
Db 3 GR 4

RESULT 8  
PT0553  
A:Title: T-cell receptor beta chain V-D-J region (126-1C) - mouse (fragment)  
C:Species: Mus musculus (house mouse)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:57:31 ; Search time 12.53 seconds  
(without alignments)  
42.556 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_163\_169

Perfect score: 7

Sequence: 1 PYVGRK 7

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1098

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database :

PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2	28.6	4	1 ECXAA	antho-RFamide neur
2	2	28.6	4	2 I40870	phospholipase C (E
3	2	28.6	4	2 A25844	antho-RF amide neu
4	2	28.6	5	2 I40702	primase - Citrobac
5	2	28.6	5	2 P00009	angiotensin-conver
6	2	28.6	5	2 S53595	hypothetical prote
7	2	28.6	5	2 PT0608	T-cell receptor be
8	2	28.6	5	2 PT0553	T-cell receptor be
9	2	28.6	6	2 A35890	RNA-directed DNA p
10	2	28.6	6	2 S11024	hydrogensulfite re
11	2	28.6	6	2 A60494	antineoplastic gly
12	2	28.6	6	2 I51434	H4 histone - Afric
13	2	28.6	6	2 I48126	alpha-tubulin - Ch
14	2	28.6	6	2 PT0511	T-cell receptor be
15	2	28.6	6	2 PT0516	T-cell receptor be
16	2	28.6	7	2 S16364	opacity protein P.
17	2	28.6	7	2 S16365	opacity protein P.
18	2	28.6	7	2 S55548	mcrB protein - Esc
19	2	28.6	7	2 E33932	Ig mu chain D regi
20	2	28.6	7	2 PT0667	T-cell receptor be
21	2	28.6	7	2 PT0655	T-cell receptor be
22	2	28.6	7	2 A38671	peptidylglycine mo
23	2	28.6	7	2 A38081	amine oxidase (cop
24	2	28.6	8	2 S37141	rpsA protein - Erw
25	2	28.6	8	2 A25836	L-serine dehydrata
26	2	28.6	8	2 S63493	dissimilatory sulf
27	2	28.6	8	2 T48890	hypothetical prote
28	2	28.6	8	2 PT0030	inulinase (EC 3.2.
29	2	28.6	8	2 D47393	neuropeptide calla

#### ALIGNMENTS

RESULT 1

ECXAA

antho-RFamide neuropeptide - sea anemone (Anthopleura elegantissima)

C:Species: Anthopleura elegantissima

C>Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 08-Dec-1995

C:Accession: A26666

R:Grimmelikhuijzen, C.J.P.; Graff, D.

Proc. Natl. Acad. Sci. U.S.A. 83, 9817-9821, 1986

A:Title: Isolation of <Glu-Gly-Arg-Phe-NH2 (Antho-RFamide), a neuropeptide from sea a

A:Reference number: A26666; MUID:87092339

A:Accession: A26666

A:Molecule type: protein

A:Residues: 1-4 <GR>

C:Comment: The function of this peptide is not known but it could act as a transmitt

C:Superfamily: RFamide neuropeptide

C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:4/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 28.6%; Score 2; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GR 6

DB 2 GR 3

RESULT 2

I40870

phospholipase C (EC 3.1.4.3) - Clostridium perfringens (fragment)

C:Species: Clostridium perfringens

C>Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 21-Jul-2000

C:Accession: I40870

R:Toyonaga, T.; Matsushita, O.; Katayama, S.; Minami, J.; Okabe, A.

Microbiol. Immunol. 36, 603-613, 1992

A:Title: Role of the upstream region containing an intrinsic DNA curvature in the neg

A:Reference number: I40870; MUID:92396045

A:Accession: I40870

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-4 <RES>

A:Cross-references: EMBL:X62825; NID:g40622; PIDN:CAA44636.1; PID:g4377417

C:Genetics:

C:Gene: plc

C:Keywords: phosphoric diester hydrolase

Query Match 28.6%; Score 2; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;



; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-755-496A-36

Query Match 57.1%; Score 4; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 6  
|  
|  
|  
|  
Db 6 VKGR 9

RESULT 15  
US-08-843-035-2  
; Sequence 2, Application US/08843035  
; Patent No. 5783662  
; GENERAL INFORMATION:  
; APPLICANT: Janney, Paul A.  
; APPLICANT: Cunningham, C. Casey  
; APPLICANT: Hartwig, John H.  
; APPLICANT: Stossel, Thomas P.  
; APPLICANT: Vegner, Roland  
; TITLE OF INVENTION: POLYPHOSPHOINOSITIDE BINDING  
; TITLE OF INVENTION: PEPTIDES FOR INTRACELLULAR DRUG DELIVERY  
; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: Federal Reserve Plaza, 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02210

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/843,035  
; FILING DATE: 11-APR-1997  
; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/394,027  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7029

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-843-035-2

Query Match 57.1%; Score 4; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 6  
|  
|  
|  
|  
Db 6 VKGR 9

Search completed: September 15, 2001, 12:58:55  
Job time: 109 sec

;; TITLE OF INVENTION: PEPTIDES FOR INTRACELLULAR DRUG DELIVERY  
;; NUMBER OF SEQUENCES: 36  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
;; STREET: Federal Reserve Plaza, 600 Atlantic Avenue  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02210  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/843,035  
;; FILING DATE: 11-APR-1997  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/394,027  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Plumer, Elizabeth R.  
;; REGISTRATION NUMBER: 36,637  
;; REFERENCE/DOCKET NUMBER: B0801/7029  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-720-3500  
;; TELEFAX: 617-720-2441  
;; INFORMATION FOR SEQ ID NO: 36:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal  
;; US-08-843-035-36

Query Match 57.1%; Score 4; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 6  
|||  
Db 6 VKGR 9

RESULT 13  
US-08-755-496A-13  
; Sequence 13, Application US/08755496A  
; Patent No. 5846743  
; GENERAL INFORMATION:  
; APPLICANT: Cunningham, Casey  
; APPLICANT: Hartwig, John H.  
; APPLICANT: Janney, Paul A.  
; APPLICANT: Stossel, Thomas P.  
; APPLICANT: Vegners, Roland  
; TITLE OF INVENTION: POLYPHOSPHONOSITIDE-(PPI-) BINDING  
; TITLE OF INVENTION: PEPTIDE(S) FOR INTRACELLULAR DRUG DELIVERY  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: Federal Reserve Plaza, 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/755,496A  
;; FILING DATE: 22-NOV-1996  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/394,027  
;; FILING DATE: 22-FEB-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Plumer, Elizabeth R.  
;; REGISTRATION NUMBER: 36,637  
;; REFERENCE/DOCKET NUMBER: B0801/7062  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-720-3500  
;; TELEFAX: 617-720-2441  
;; INFORMATION FOR SEQ ID NO: 13:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 9 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal  
;; US-08-755-496A-13

Query Match 57.1%; Score 4; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 6  
|||  
Db 5 VKGR 8

RESULT 14  
US-08-755-496A-36  
; Sequence 36, Application US/08755496A  
; Patent No. 5846743  
; GENERAL INFORMATION:  
; APPLICANT: Cunningham, Casey  
; APPLICANT: Hartwig, John H.  
; APPLICANT: Janney, Paul A.  
; APPLICANT: Stossel, Thomas P.  
; APPLICANT: Vegners, Roland  
; TITLE OF INVENTION: POLYPHOSPHONOSITIDE-(PPI-) BINDING  
; TITLE OF INVENTION: PEPTIDE(S) FOR INTRACELLULAR DRUG DELIVERY  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: Federal Reserve Plaza, 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/755,496A  
; FILING DATE: 22-NOV-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/394,027  
; FILING DATE: 22-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7062  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid

REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7062  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-755-496A-12

Query Match 57.1%; Score 4; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 6  
DB 4 VKGR 7

RESULT 10  
US-08-030-077-4  
Sequence 4, Application US/08030077  
Patent No. 5451658  
GENERAL INFORMATION:  
APPLICANT: Seelig, Gail F.  
TITLE OF INVENTION: Antagonists of Human Gamma Interferon  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schering-Plough Corporation  
STREET: One Giralda Farms  
CITY: Madison  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07940  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 6.0.5  
SOFTWARE: Microsoft Word 4.00B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/030,077  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/030,077  
FILING DATE: 19-March-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lunn, Paul G.  
REGISTRATION NUMBER: 32,743  
REFERENCE/DOCKET NUMBER: JB0151K  
TELEPHONE: 201-822-7255  
TELEFAX: 201-822-7039  
TELEX: 219165  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-030-077-4

Query Match 57.1%; Score 4; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYVK 4

DB 6 PYVK 9

RESULT 11  
US-08-843-035-13  
Sequence 13, Application US/08843035  
Patent No. 5783662  
GENERAL INFORMATION:  
APPLICANT: Janney, Paul A.  
APPLICANT: Cunningham, C. Casey  
APPLICANT: Hartwig, John H.  
APPLICANT: Stossel, Thomas P.  
APPLICANT: Vegner, Roland  
TITLE OF INVENTION: POLYPHOSPHONOSITIDE BINDING  
TITLE OF INVENTION: PEPTIDES FOR INTRACELLULAR DRUG DELIVERY  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: Federal Reserve Plaza, 500 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/843,035  
FILING DATE: 11-APR-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/394,027  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7029  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-843-035-13

Query Match 57.1%; Score 4; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 6  
DB 5 VKGR 8

RESULT 12  
US-08-843-035-36  
Sequence 36, Application US/08843035  
Patent No. 5783662  
GENERAL INFORMATION:  
APPLICANT: Janney, Paul A.  
APPLICANT: Cunningham, C. Casey  
APPLICANT: Hartwig, John H.  
APPLICANT: Stossel, Thomas P.  
APPLICANT: Vegner, Roland  
TITLE OF INVENTION: POLYPHOSPHONOSITIDE BINDING



; Sequence 11, Application US/08755496A  
; Patent No. 5846743

## ; GENERAL INFORMATION:

; APPLICANT: Cunningham, Casey  
; APPLICANT: Hartwig, John H.  
; APPLICANT: Janney, Paul A.  
; APPLICANT: Stossel, Thomas P.  
; APPLICANT: Vegners, Roland

; TITLE OF INVENTION: POLYPHOSPHINOSITIDE-(PPI-) BINDING  
; TITLE OF INVENTION: PEPTIDE(S) FOR INTRACELLULAR DRUG DELIVERY  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: Federal Reserve Plaza, 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA

; COUNTRY: USA

; ZIP: 02210

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08755.496A  
; FILING DATE: 22-NOV-1996

## ; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/394.027  
; FILING DATE: 22-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7062  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

US-08-755-496A-11

Query Match 57.1%; Score 4; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 6

|||||

Db 3 VKGR 6

## RESULT 8

US-08-843-035-12  
; Sequence 12, Application US/08843035  
; Patent No. 5783662

## ; GENERAL INFORMATION:

; APPLICANT: Janney, Paul A.  
; APPLICANT: Cunningham, C. Casey  
; APPLICANT: Hartwig, John H.  
; APPLICANT: Stossel, Thomas P.  
; APPLICANT: Vegner, Roland

; TITLE OF INVENTION: POLYPHOSPHINOSITIDE BINDING

; TITLE OF INVENTION: PEPTIDES FOR INTRACELLULAR DRUG DELIVERY

; NUMBER OF SEQUENCES: 36

## ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wolf, Greenfield &amp; Sacks, P.C.

; STREET: Federal Reserve Plaza, 600 Atlantic Avenue

; CITY: Boston

; STATE: MA

; COUNTRY: USA  
; ZIP: 02210

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/843.035  
; FILING DATE: 11-APR-1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/394.027  
; FILING DATE:

## ; ATTORNEY/AGENT INFORMATION:

; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7029  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

US-08-843-035-12

## Query Match

57.1%; Score 4; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 6

|||||

Db 4 VKGR 7

## RESULT 9

US-08-755-496A-12  
; Sequence 12, Application US/08755496A  
; Patent No. 5846743

## ; GENERAL INFORMATION:

; APPLICANT: Cunningham, Casey  
; APPLICANT: Hartwig, John H.  
; APPLICANT: Janney, Paul A.  
; APPLICANT: Stossel, Thomas P.  
; APPLICANT: Vegners, Roland

; TITLE OF INVENTION: POLYPHOSPHINOSITIDE-(PPI-) BINDING

; TITLE OF INVENTION: PEPTIDE(S) FOR INTRACELLULAR DRUG DELIVERY

; NUMBER OF SEQUENCES: 38

## ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wolf, Greenfield &amp; Sacks, P.C.

; STREET: Federal Reserve Plaza, 600 Atlantic Avenue

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02210

## ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/755.496A  
; FILING DATE: 22-NOV-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/394.027  
; FILING DATE: 22-FEB-1995

## ; ATTORNEY/AGENT INFORMATION:

; NAME: Plumer, Elizabeth R.

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/394,027  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7029  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-843-035-10

Query Match 57.1%; Score 4; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 6  
DB 2 VKGR 5

RESULT 5  
US-08-755-496A-10  
Sequence 10, Application US/08755496A  
Patent No. 5846743  
GENERAL INFORMATION:  
APPLICANT: Cunningham, Casey  
APPLICANT: Hartwig, John H.  
APPLICANT: Janney, Paul A.  
APPLICANT: Stossel, Thomas P.  
APPLICANT: Vegners, Roland  
TITLE OF INVENTION: POLYPHOSPHOINOSITIDE-(PPI-) BINDING  
TITLE OF INVENTION: PEPTIDE(S) FOR INTRACELLULAR DRUG DELIVERY  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: Federal Reserve Plaza, 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/755,496A  
FILING DATE: 22-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/394,027  
FILING DATE: 22-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7062  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

FRAGMENT TYPE: internal  
US-08-755-496A-10

Query Match 57.1%; Score 4; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 6  
DB 2 VKGR 5

RESULT 6  
US-08-843-035-11  
Sequence 11, Application US/08843035  
Patent No. 5783662  
GENERAL INFORMATION:  
APPLICANT: Janney, Paul A.  
APPLICANT: Cunningham, C. Casey  
APPLICANT: Hartwig, John H.  
APPLICANT: Stossel, Thomas P.  
APPLICANT: Vegner, Roland  
TITLE OF INVENTION: POLYPHOSPHOINOSITIDE BINDING  
TITLE OF INVENTION: PEPTIDES FOR INTRACELLULAR DRUG DELIVERY  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: Federal Reserve Plaza, 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/843,035  
FILING DATE: 11-APR-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/394,027  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7029  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-843-035-11

Query Match 57.1%; Score 4; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 6  
DB 3 VKGR 6

RESULT 7  
US-08-755-496A-11

```
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VKGR 6
Db 1 VKGR 4

RESULT 2
US-08-755-496A-9
; Sequence 9, Application US/08755496A
; Patent No. 5846743
; GENERAL INFORMATION:
; APPLICANT: Cunningham, Casey
; APPLICANT: Hartwig, John H.
; APPLICANT: Janney, Paul A.
; APPLICANT: Stossel, Thomas P.
; APPLICANT: Vegners, Roland
; TITLE OF INVENTION: POLYPHOSPHONOSITIDE-(PPI-) BINDING
; TITLE OF INVENTION: PEPTIDE(S) FOR INTRACELLULAR DRUG DELIVERY
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: Federal Reserve Plaza, 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 22-FEB-1995
; APPLICATION NUMBER: 08/394,027
; FILING DATE: 22-NOV-1996
; APPLICATION NUMBER: US/08/755,496A
; PRIOR APPLICATION DATA:
; FILING DATE: 22-FEB-1995
; APPLICATION NUMBER: 08/394,027
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-755-496A-9

Query Match 57.1%; Score 4; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VKGR 6
Db 1 VKGR 4

RESULT 3
US-08-812-586-20
; Sequence 20, Application US/08812586
; Patent No. 6048704
; GENERAL INFORMATION:
; APPLICANT: Martin David Tilson
; TITLE OF INVENTION: PURIFIED AND RECOMBINANT ANTIGENIC
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH ABDOMINAL AORTIC ANEURYSM (AAA)

; TITLE OF INVENTION: DISEASE, AND DIAGNOSTIC AND THERAPEUTIC USES THEREOF
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 07-NAR-1997
; APPLICATION NUMBER: US/08/812,586
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/53862-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-812-586-20

Query Match 57.1%; Score 4; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KGRK 7
Db 1 KGRK 4

RESULT 4
US-08-843-035-10
; Sequence 10, Application US/08843035
; Patent No. 5783662
; GENERAL INFORMATION:
; APPLICANT: Janney, Paul A.
; APPLICANT: Cunningham, C. Casey
; APPLICANT: Hartwig, John H.
; APPLICANT: Stossel, Thomas P.
; APPLICANT: Vegner, Roland
; TITLE OF INVENTION: POLYPHOSPHONOSITIDE BINDING
; TITLE OF INVENTION: PEPTIDES FOR INTRACELLULAR DRUG DELIVERY
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: Federal Reserve Plaza, 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 11-APR-1997
; APPLICATION NUMBER: US/08/843,035
; CLASSIFICATION: 530
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:57:06 ; Search time 12.39 Seconds  
(without alignments)  
11.633 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_163\_169

Perfect score: 7

Sequence: 1 PVVKGK 7

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 59481

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgml\_7/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgml\_7/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgml\_7/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgml\_7/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgml\_7/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgml\_7/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4	57.1	5	1	US-08-843-035-9
2	4	57.1	5	2	US-08-755-496A-9
3	4	57.1	5	3	US-08-812-586-20
4	4	57.1	6	1	US-08-843-035-10
5	4	57.1	6	2	US-08-755-496A-10
6	4	57.1	7	1	US-08-843-035-11
7	4	57.1	7	2	US-08-755-496A-11
8	4	57.1	8	1	US-08-843-035-12
9	4	57.1	8	2	US-08-755-496A-12
10	4	57.1	9	1	US-08-030-077-4
11	4	57.1	9	1	US-08-843-035-13
12	4	57.1	9	2	US-08-843-035-36
13	4	57.1	9	2	US-08-755-496A-13
14	4	57.1	9	2	US-08-755-496A-36
15	4	57.1	10	1	US-08-843-035-2
16	4	57.1	10	1	US-08-843-035-17
17	4	57.1	10	2	US-08-248-839C-78
18	4	57.1	10	2	US-08-755-496A-2
19	4	57.1	10	2	US-08-755-496A-17
20	4	57.1	10	2	US-08-685-589A-127
21	3	42.9	3	3	US-08-405-647B-3
22	3	42.9	3	4	US-08-985-499-3
23	3	42.9	3	5	PCT-US96-03180-3
24	3	42.9	4	1	US-07-945-982-5
25	3	42.9	4	1	US-08-082-847-1
26	3	42.9	4	1	US-08-372-455-5
27	3	42.9	4	1	US-08-286-888B-9

28	3	42.9	4	1	US-08-053-131-5	Sequence 5, Appli
29	3	42.9	4	1	US-08-430-536A-22	Sequence 22, Appl
30	3	42.9	4	1	US-08-645-641-5	Sequence 5, Appli
31	3	42.9	4	1	US-08-361-862-3	Sequence 3, Appli
32	3	42.9	4	1	US-08-843-035-8	Sequence 8, Appli
33	3	42.9	4	1	US-08-638-271A-3	Sequence 3, Appli
34	3	42.9	4	1	US-07-853-408B-5	Sequence 5, Appli
35	3	42.9	4	2	US-08-096-762-5	Sequence 5, Appli
36	3	42.9	4	2	US-08-755-496A-8	Sequence 8, Appli
37	3	42.9	4	2	US-08-308-865-5	Sequence 5, Appli
38	3	42.9	4	2	US-08-685-589A-47	Sequence 47, Appl
39	3	42.9	4	2	US-08-394-189B-19	Sequence 19, Appl
40	3	42.9	4	2	US-08-684-547-22	Sequence 22, Appl
41	3	42.9	4	4	US-08-859-242-42	Sequence 42, Appl
42	3	42.9	4	5	PCT-US92-10983-5	Sequence 5, Appli
43	3	42.9	4	5	PCT-US93-08479-2	Sequence 2, Appli
44	3	42.9	5	1	US-08-325-509-44	Sequence 44, Appl
45	3	42.9	5	1	US-08-097-938-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1  
US-08-843-035-9  
; Sequence 9, Application US/08843035  
; Patent No. 5783862  
; GENERAL INFORMATION:  
; APPLICANT: Jannney, Paul A.  
; APPLICANT: Cunningham, C. Casey  
; APPLICANT: Hartwig, John H.  
; APPLICANT: Stossel, Thomas P.  
; APPLICANT: Wegner, Roland  
; TITLE OF INVENTION: POLYPHOSPHONOSITIDE BINDING  
; TITLE OF INVENTION: PEPTIDES FOR INTRACELLULAR DRUG DELIVERY  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: Federal Reserve Plaza, 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/843,035  
; FILING DATE: 11-APR-1997  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/394,027  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plummer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7029  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-843-035-9

Query Match 57.1%; Score 4; DB 1; Length 5;

**THIS PAGE BLANK (USPTO)**





PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
PI Cunningham C, Hartwig JH, Janney PA, Stossel TP;  
PR Vegner R;  
XX  
XX WPI: 1998-297605/26.  
XX  
XX Carrier molecule for transporting membrane-impermeable agents across  
PT cell membranes - comprises N-terminal derivatised peptide that binds  
PT to poly:phospho:inositol, used for delivering, e.g. antibiotics and  
PT nucleic acids for therapy  
XX  
XX  
XX  
PS Claim 7; Page 9; 70pp; English.  
XX  
XX Sequences shown in AAM61920 to AAM61925 and AAM61942 to AAM61948 are  
CC peptide fragments of the transport mediating peptide P2 (AAM61914) which  
CC can bind to a polyphosphoinositide (PPI). Peptides shown in AAM61913 (P1)  
CC and AAM61914 (P2) correspond to the amino acid sequences of PPI-binding  
CC domains of human gelsolin. Fragments and functional equivalents of these  
CC peptides shown in sequences AAM61915 to AAM61948 are also capable of  
CC binding to PPI. These PPI-binding peptides can be used in a carrier  
CC molecule along with a covalently coupled N-terminal derivatising agent  
CC for facilitating transport of a membrane-impermeable extracellular agent,  
CC having an intracellular activity, across a cell membrane. The carrier  
CC molecules are coupled to a very wide range of membrane-impermeable  
CC extracellular agent, most preferably an antibiotic, oligonucleotide or  
CC imaging agent, to form a prodrug that delivers the agent into a cell,  
CC particularly within 10-300 seconds. The process may be carried out in  
CC vitro or in vivo and typical applications include, eliminating  
CC intracellular bacteria from cell cultures, transfection with nucleic  
CC acid, including antisense molecules, delivering probes and primers for in  
CC situ hybridisation or amplification, delivery of therapeutic drugs and  
CC also modulation of PPI-mediated signalling, e.g. for treating psoriasis  
CC or preventing post-angioplasty thrombosis or restenosis. The carrier  
CC molecules can also be used to prevent cold-induced platelet activation by  
CC binding to PPI generated during activation, preventing further  
CC interaction with actin-filament capping proteins. They provide rapid and  
CC efficient delivery of the extracellular agent to a wide variety of cells,  
CC even where these are fixed or are at low temperature (4 deg. C). The  
CC N-terminal derivatising agent not only blocks the N-terminus but can also  
CC serve as detectable label. The carrier blocks the N-terminus but can also  
CC with PPI on the cell membrane, so deliver the extracellular agent by a  
CC mechanism that is independent of receptor-mediated endocytosis and active  
CC cellular metabolism.  
XX  
XX Sequence 9 AA:  
SQ

Query Match 57.1%; Score 4; DB 19; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 3 VKGR 6  
| | | |  
Qb 5 VKGR 8

RESULT 13  
AAM61948  
ID AAM61948 standard; peptide; 9 AA.  
XX  
XX AAM61948;  
XX  
XX 17-SEP-1998 (first entry)  
XX  
XX PPI binding transport mediating peptide P2 fragment 13.  
XX  
XX Polyphosphoinositide; transport; impermeable agent; PPI; human;  
KW antibiotic; psoriasis; intracellular drug delivery; restenosis;  
KW post-angioplasty thrombosis; PPI-mediated signalling; gelsolin;  
KW phosphoinositol; platelet activation.  
XX  
XX Synthetic.  
SQ

OS Homo sapiens.  
XX  
XX WO9820867-A1.  
XX PD  
XX 22-MAY-1998.  
XX PF 14-NOV-1996; 96WO-US18453.  
XX PR 14-NOV-1996; 96WO-US18453.  
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
PI Cunningham C, Hartwig JH, Janney PA, Stossel TP;  
PI Vegner R;  
DR WPI: 1998-297605/26.  
XX  
XX Carrier molecule for transporting membrane-impermeable agents across  
PT cell membranes - comprises N-terminal derivatised peptide that binds  
PT to polyphosphoinositol, used for delivering, e.g. antibiotics and  
PT nucleic acids for therapy  
XX  
PS Claim 7; Page 9; 70pp; English.

Sequences shown in AAW61920 to AAW61925 and AAW61942 to AAW61948 are peptide fragments of the transport mediating peptide p2 (AAW61914) which can bind to a polyphosphoinositide (PPI). Peptides shown in AAW61913 (P1) and AAW61914 (P2) correspond to the amino acid sequences of PPI-binding domains of human gelsolin. Fragments and functional equivalents of these peptides shown in sequences AAW61915 to AAW61948 are also capable of binding to PPI. These PPI-binding peptides can be used in a carrier molecule along with a covalently coupled N-terminal derivatising agent for facilitating transport of a membrane-impermeable extracellular agent, having an intracellular activity, across a cell membrane. The carrier molecules are coupled to a very wide range of membrane-impermeable extracellular agent, most preferably an antibiotic, oligonucleotide or imaging agent, to form a product that delivers the agent into a cell, particularly within 10-300 seconds. The process may be carried out in vitro or in vivo and typical applications include, eliminating intracellular bacteria from cell cultures, transfection with nucleic acid, including antisense molecules, delivering probes and primers for in situ hybridisation or amplification, delivery of therapeutic drugs and also modulation of PPI-mediated signalling, e.g. for treating psoriasis or preventing post-angioplasty thrombosis or restenosis. The carrier molecules can also be used to prevent cold-induced platelet activation by binding to PPI generated during activation, preventing further interaction with actin-filament capping proteins. They provide rapid and efficient delivery of the extracellular agent to a wide variety of cells, even where these are fixed or are at low temperature (4 deg. C). The N-terminal derivatising agent not only blocks the N-terminus but can also serve as detectable label. The carrier molecules are believed to interact with PPI on the cell membrane, so deliver the extracellular agent by a mechanism that is independent of receptor-mediated endocytosis and active cellular metabolism.

SQ Sequence 9 AA:

Query Match 57.1%; Score 4; DB 19; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3,4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 3 VKGR 6  
| | |  
DB 6 VKGR 9

RESULT 14  
AAR31033  
ID AAR31033 standard; Protein; 10 AA.  
AC  
XX AAR31033;  
XX



OY 1 PYVK 4  
 ||||  
 DB 3 pyvk 6

## RESULT 10

AAR23129  
 ID AAR23129 standard; peptide; 9 AA.

AC AAR23129;

DF 26-OCT-1992 (first entry)

DE Synthetic gamma Interferon antagonist.

KM IFN $\gamma$  antibodies; anti-idiotypic; antagonist; autoimmune; multiple;  
 KM sclerosis; screening.

OS Homo sapiens.

PN WO9206115-A.

PD 16-APR-1992.

PF 25-SEP-1991; 91WO-US06771.

PR 27-SEP-1990; 90US-0589106.

PA (SCHE ) SCHERING CORP.

PI Seelig G;

DR WPI; 1992-150818/18.

XX Polypeptide(s) antibodies and anti idiotypic antibodies - used as  
 PT antagonists of human gamma interferon for treatment of e.g.  
 PT autoimmune disease

PS Claim 2; Page 37; 45pp; English.

CC Novel polypeptides contg. up to 50 amino acids comprising one or more  
 CC amino acid subsequences selected from residues 15-21 and 132-137 of  
 CC the sequence shown in AAR23126 may be used as antagonists of human  
 CC gamma IFN, and may also be used to raise antibodies or anti-idiotypic  
 CC antibodies against human gamma IFN. The peptides and antibodies  
 CC were used to treat diseases mediated by IFN-gamma e.g., autoimmune  
 CC diseases and multiple sclerosis, and for in vitro study of the  
 CC mechanism of IFN-gamma binding to various cell types and for  
 CC screening for other IFN-gamma antagonists or agonists.  
 CC See also AAR23126-35.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 13; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PYVK 4  
 ||||  
 DB 6 pyvk 9

## RESULT 11

AAR80484  
 ID AAR80484 standard; peptide; 9 AA.

AC AAR80484;

DF 05-MAR-1996 (first entry)

XX Recombinant bovine adseverin peptide fragment residues 138-146.  
 DE

XX Bovine adseverin; actin filament; cleavage; calcium ion modulated;  
 KM antithrombotic agent; recombinant; fragment; residues 138-146.  
 XX

OS Bos taurus.

PN WO9518221-A1.

PD 06-JUL-1995.

PF 27-DEC-1994; 94WO-JP02227.

PR 20-DEC-1994; 94JP-0340692.

PR 28-DEC-1993; 93JP-0355112.

PR 12-JUL-1994; 94JP-0160236.

PA (CHUS ) CHUGAI SEIYAKU KK.

PA (NAKA/) NAKAMURA N.

PI Nakamura S, Nezu J, Sakurai T;

DR WPI; 1995-246387/32.

XX Novel DNA encoding a recombinant adseverin and also anti-sense  
 PT oligo:nucleotide(s) and antibodies - involved in calcium  
 PT ion-modulated cleavage of actin filaments and is useful as an  
 PT antithrombotic agent.

PS Disclosure; Page 10; 64pp; Japanese.

CC AAR80483 and AAR80484 are recombinant bovine adseverin (RBA) peptide  
 CC fragments. The RBA has similar activity to natural adseverin in the  
 CC calcium ion modulated cleavage of actin filaments, and can therefore  
 CC be used as an antithrombotic agent.

XX Sequence 9 AA;

Query Match 57.1%; Score 4; DB 16; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 VKGR 6  
 ||||  
 DB 5 vkgr 8

## RESULT 12

AAM61925  
 ID AAM61925 standard; peptide; 9 AA.

AC AAM61925;

DF 17-SEP-1998 (first entry)

DE PPI binding transport mediating peptide P2 fragment 6.

XX Polypeptide(s) transport; impermeable agent; PPI; human;  
 KM antihypertensive; psoriasis; intracellular drug delivery; restenosis;  
 KM post-angioplasty thrombosis; PPI-mediated signalling; gelatin;  
 KM phosphoinositol; platelet activation.

OS Synthetic.

OS Homo sapiens.

PN WO9820887-A1.

PD 22-MAY-1998.

PF 14-NOV-1996; 96WO-US18453.

PR 14-NOV-1996; 96WO-US18453.

```
XX WPI; 2000-376575/32.
DR
XX
XX Preparation of peptide with C-terminal alpha-carboxamide residue, e.g.
PT growth hormone releasing factors comprises treating substrate with
PT ammonia in presence of clostripain
XX
XX Example 7; Page 19; 48pp; English.
PS
XX
XX The present sequence is a synthetic peptide used in enzymatic amidation.
CC This peptide has a genetic structure and residues 5 to 8 may be any of
CC the peptides described in AAB07300 to AAB07311. The present sequence was
CC used to produce a peptide with a C-terminal alpha-carboxamide (AAB07299)
CC by treating the present peptide with an ammonia reagent and clostripain
CC (also known as clostridopeptidase B). Clostripain is an extracellular
CC thiol endoprotease from Clostridia. Clostripain cleaves arginine
CC containing peptides amidatively at an Arg-Xaa peptide bond.
XX
SQ Sequence 8 AA;

Query Match
Best Local Similarity 57.1%; Score 4; DB 21; Length 8;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 VKGR 6
DB 1 VKGR 4

RESULT 8
AAP40533
ID AAP40533 standard; peptide; 9 AA.
XX
XX AAP40533;
AC
XX
XX 24-JAN-1992 (first entry)
DT
XX
XX Sequence of a portion of human interferon (IFN) gamma.
DE
XX
XX Killer cell cytotoxicity; virucide; interferon gamma; antitumour;
KM immunostimulant.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /label= N-Ac-Gln
XX
XX US4473555-A.
XX
XX 25-SEP-1984.
XX
XX 17-OCT-1983; 83US-0542632.
XX
XX 17-OCT-1983; 83US-0542632.
XX
XX 17-OCT-1983; 83US-0542632.
XX
XX (SYNT ) SYNTEX (USA) INC.
XX
XX Nestor JJ, Mofatt JG, Slims JM;
XX
XX WPI; 1984-256353/41.
XX
XX Claim 2; column 8; 5pp; English.
XX
XX The peptides of the invention augment natural killer cell
CC cytotoxicity and comprise in whole or in part the N-terminus of a
CC polypeptide which is a transcription of a DNA fragment determined to
CC be the producer of the protein portion of human interferon gamma.
CC Also possible is the use of recombinant DNA technology, see e.g.
CC EP--32134 (81-53697 D/30) or EP--48970 (82-2897E/15). A typical
CC pharmaceutical compsn. comprises 10.0mg polypeptide, 86.0mg
CC compressible sugar USP, 4.0mg Ca stearate, in a form suitable for
CC sublingual admin.
XX
XX Sequence 9 AA;
```

```
CC Also possible is the use of recombinant DNA technology, see e.g.
CC EP--32134 (81-53697 D/30) or EP--48970 (82-2897E/15). A typical
CC pharmaceutical compsn. comprises 10.0mg polypeptide, 86.0mg
CC compressible sugar USP, 4.0mg Ca stearate, in a form suitable for
CC sublingual admin.
XX
XX Sequence 9 AA;

Query Match
Best Local Similarity 57.1%; Score 4; DB 5; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PYVK 4
DB 3 PYVK 6

RESULT 9
AAP40534
ID AAP40534 standard; peptide; 9 AA.
XX
XX AAP40534;
AC
XX
XX 24-JAN-1992 (first entry)
DT
XX
XX Sequence of a portion of human interferon (IFN) gamma.
DE
XX
XX Killer cell cytotoxicity; virucide; interferon gamma; antitumour;
KM immunostimulant.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /label= N-Ac-Gln
XX
XX US4473555-A.
XX
XX 25-SEP-1984.
XX
XX 17-OCT-1983; 83US-0542632.
XX
XX 17-OCT-1983; 83US-0542632.
XX
XX 17-OCT-1983; 83US-0542632.
XX
XX (SYNT ) SYNTEX (USA) INC.
XX
XX Nestor JJ, Mofatt JG, Slims JM;
XX
XX WPI; 1984-256353/41.
XX
XX Claim 3; column 8; 5pp; English.
XX
XX The peptides of the invention augment natural killer cell
CC cytotoxicity and comprise in whole or in part the N-terminus of a
CC polypeptide which is a transcription of a DNA fragment determined to
CC be the producer of the protein portion of human interferon gamma.
CC Prepn. is by usual synthetic methods, partic. solid phase synthesis.
CC Also possible is the use of recombinant DNA technology, see e.g.
CC EP--32134 (81-53697 D/30) or EP--48970 (82-2897E/15). A typical
CC pharmaceutical compsn. comprises 10.0mg polypeptide, 86.0mg
CC compressible sugar USP, 4.0mg Ca stearate, in a form suitable for
CC sublingual admin.
XX
XX Sequence 9 AA;
```



CC approximately 40KD. The protein is isolated from human aortic tissue and  
CC binds immunoreactively with immunoglobulin purified from human abdominal  
CC aortic aneurysm (AAA) tissue. The protein is referred to as aortic  
CC aneurysm-associated antigenic protein (AAP). The protein is capable of  
CC forming a disulphide bonded dimer. The protein is immunoreactive with  
CC human kappa immunoglobulin. Also included in the invention are  
CC recombinantly produced human AAA proteins. AAP shows regions of homology  
CC with the bovine microfibril associated glycoprotein MFAP-4 and also with  
CC fibrinogen and vitronectin. The isolated microfibrillar protein is useful  
CC for alleviating abdominal aortic aneurysm (AAA) disease and detecting the  
CC presence of AAA-associated immunoglobulin bound to the human aortic  
CC tissue. Antibodies directed against AAP can be used to detect AAA  
CC disease. The recombinant protein can be used to induce tolerance to  
CC antigenic AAA protein in the subject e.g. human. This sequence represents  
CC a calcium binding motif, this sequence is used to identify and  
CC characterise the AAP of the invention.

SO Sequence 5 AA;

Query Match 57.1%; Score 4; DB 21; Length 5;

Best Local Similarity 100.0%; Pred. No. 3.4e+05; Mismatches 0; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 KGRK 7  
|||  
Db 1 kgfk 4

RESULT 4

AAW18608 standard; peptide; 6 AA.

AC AAW18608;

DT 04-MAR-1998 (first entry)

DE Aged band 3 peptide (residues 822-827) epitope.

KW Band 3 protein; antibody; aging antigenic site; Alzheimer's disease;  
phosphorylation; detection; epitope.

OS Homo sapiens.

PN WO9726537-A1.

PD 24-JUL-1997.

PF 13-DEC-1996; 96WO-US20465.

PR 19-JAN-1996; 96US-0010250.

PA (RESE ) RESEARCH CORP TECHNOLOGIES INC.

PI Kay MMB;

DR WPI; 1997-385478/35.

PT Detecting Alzheimer's disease using antibody that recognises aged  
band 3 protein in tissues - or from reduced degree of band 3 protein  
phosphorylation, can be applied to blood or brain samples

PS Disclosure; Page 12; 45pp; English.

CC This is an aging antigenic band 3 peptide (residues 822-827) to which a  
CC specific antibody can bind to. Band 3 is a ubiquitous anion-exchange  
CC protein and ages as cells and tissues age. Antibodies have been developed  
CC against this aged band 3. These antibodies bind to distinct regions of  
CC band 3 in old cells (aging antigenic sites) but not middle aged or young  
CC cells. This can be used for detecting Alzheimer's disease. A tissue  
CC sample containing band 3 from a patient suspected of having Alzheimer's  
CC disease is treated with an antibody that can differentiate between the  
CC Alzheimer's (aged) and normal band 3, under complex-forming conditions

CC and detecting any complex formed. A tissue containing band 3 from a  
CC healthy control is treated in a similar manner and the amounts of complex  
CC formed are compared. A significantly greater formation of complex in the  
CC suspect sample as compared with that of the control is indicative of  
CC Alzheimer's disease. The disease can also be detected by comparing the  
CC degrees of phosphorylation of band 3 or its degradation products in  
CC suspect and control samples. A significant decrease in phosphorylation  
CC in the suspect sample indicates Alzheimer's disease.

SO Sequence 6 AA;

Query Match 57.1%; Score 4; DB 18; Length 6;

Best Local Similarity 100.0%; Pred. No. 3.4e+05; Mismatches 0; Indels 0; Gaps 0;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PYVK 4  
|||  
Db 2 pyvk 5

RESULT 5

AAW61922 standard; peptide; 6 AA.

AC AAW61922;

DT 17-SEP-1998 (first entry)

DE PPI binding transport mediating peptide p2 fragment 3.

KW Polyphosphoinositide; transport; impermeable agent; PPI; human;  
antibiotic; psoriasis; intracellular drug delivery; restenosis;  
KM post-angioplasty thrombosis; PPI-mediated signalling; gelsolin;  
KW phosphoinositol; platelet activation.

OS Synthetic.

OS Homo sapiens.

PN WO9820887-A1.

PD 22-MAY-1998.

PF 14-NOV-1996; 96WO-US18453.

PR 14-NOV-1996; 96WO-US18453.

PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.

PI Cunningham C, Hartwig JH, Janney PA, Stossel TP;  
Vegner R;

DR WPI; 1998-297605/26.

PT Carrier molecule for transporting membrane-impermeable agents across  
cell membranes - comprises N-terminal derivatised peptide that binds  
to poly:phosphoinositol, used for delivering, e.g. antibiotics and  
nucleic acids for therapy

PS Claim 7; Page 9; 70pp; English.

CC Sequences shown in AAW61920 to AAW61925 and AAW61942 to AAW61948 are  
CC peptide fragments of the transport mediating peptide p2 (AAW61914) which  
CC can bind to a polyphosphoinositide (PPI). Peptides shown in AAW61913 (p1)  
CC and AAW61914 (p2) correspond to the amino acid sequences of PPI-binding  
CC domains of human gelsolin. Fragments and functional equivalents of these  
CC peptides shown in sequences AAW61915 to AAW61948 are also capable of  
CC binding to PPI. These PPI-binding peptides can be used in a carrier  
CC molecule along with a covalently coupled N-terminal derivatising agent  
CC for facilitating transport of a membrane-impermeable extracellular agent,  
CC having an intracellular activity, across a cell membrane. The carrier  
CC molecules are coupled to a very wide range of membrane-impermeable  
CC extracellular agent, most preferably an antibiotic, oligonucleotide or

PT Preparation of peptide with C-terminal alpha-carboxamide residue, e.g.  
 PT growth hormone releasing factors comprises treating substrate with  
 PT ammonia in presence of clostripain  
 XX  
 PS Example 7; Page 19; 48pp; English.  
 CC The present sequence is a synthetic peptide produced from another  
 CC synthetic peptide (AAB07298) by enzymatic amidation. The present  
 CC sequence has a C-terminal alpha-carboxamide and was produced by treating  
 CC the peptide of AAB07298 with an ammonia reagent and clostripain (also  
 CC known as clostridiopeptidase B). Clostripain is an extracellular thiol  
 CC endoprotease from Clostridia. Clostripain cleaves arginine containing  
 CC peptides amidatively at an Arg-Xaa peptide bond.  
 XX  
 SQ Sequence 4 AA;  
 QY Query Match 57.1%; Score 4; DB 21; Length 4;  
 DB Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 VKGR 6  
 DB 1 VKGR 4  
 AC AAB0721;  
 XX AAB0721;  
 DT 17-SEP-1998 (first entry)  
 DE PPI binding transport mediating peptide P2 fragment 2.  
 XX  
 XX Polyphosphoinositide; transport; impermeable agent; PPI; human;  
 KW antibiotic; psoriasis; intracellular drug delivery; restenosis;  
 KW post-angioplasty thrombosis; PPI-mediated signalling; gelsolin;  
 KW phosphoinositol; platelet activation.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN MO9820887-A1.  
 XX  
 PD 22-MAY-1998.  
 XX  
 PF 14-NOV-1996; 96MO-US18453.  
 XX  
 PR 14-NOV-1996; 96MO-US18453.  
 XX  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
 XX  
 PI Cunningham C, Hartwig JH, Janney PA, Stossel TP;  
 PI Vegner R;  
 XX  
 DR WPI: 1998-297605/26.  
 XX  
 PS Claim 7; Page 9; 70pp; English.  
 CC Sequences shown in AAB61920 to AAB61925 and AAB61942 to AAB61948 are  
 CC peptide fragments of the transport mediating peptide P2 (AAB61914) which  
 CC can bind to a polyphosphoinositide (PPI). Peptides shown in AAB61913 (P1)  
 CC and AAB61914 (P2) correspond to the amino acid sequences of PPI-binding  
 CC domains of human gelsolin. Fragments and functional equivalents of these  
 CC peptides shown in sequences AAB61915 to AAB61948 are also capable of  
 CC binding to PPI. These PPI-binding peptides can be used in a carrier

CC molecule along with a covalently coupled N-terminal derivatising agent  
 CC for facilitating transport of a membrane-impermeable extracellular agent,  
 CC having an intracellular activity, across a cell membrane. The carrier  
 CC molecules are coupled to a very wide range of membrane-impermeable  
 CC extracellular agent, most preferably an antibiotic, oligonucleotide or  
 CC imaging agent, to form a prodrug that delivers the agent into a cell,  
 CC particularly within 10-300 seconds. The process may be carried out in  
 CC vitro or in vivo and typical applications include, eliminating  
 CC intracellular bacteria from cell cultures, transfection with nucleic  
 CC acid, including antisense molecules, delivering probes and primers for in  
 CC situ hybridisation or amplification, delivery of therapeutic drugs and  
 CC also modulation of PPI-mediated signalling, e.g. for treating psoriasis  
 CC or preventing post-angioplasty thrombosis or restenosis. The carrier  
 CC molecules can also be used to prevent cold-induced platelet activation by  
 CC binding to PPI generated during activation, preventing further  
 CC interaction with actin-filament capping proteins. They provide rapid and  
 CC efficient delivery of the extracellular agent to a wide variety of cells,  
 CC even where these are fixed or are at low temperature (4 deg. C). The  
 CC N-terminal derivatising agent not only blocks the N-terminus but can also  
 CC serve as detectable label. The carrier molecules are believed to interact  
 CC with PPI on the cell membrane, so deliver the extracellular agent by a  
 CC mechanism that is independent of receptor-mediated endocytosis and active  
 CC cellular metabolism.  
 XX  
 SQ Sequence 5 AA;  
 QY Query Match 57.1%; Score 4; DB 19; Length 5;  
 DB Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 VKGR 6  
 DB 1 VKGR 4  
 AC AAB03688;  
 XX AAB03688;  
 DT 04-OCT-2000 (first entry)  
 DE Calcium binding motif 2.  
 XX  
 XX Aortic aneurysm-associated antigen protein; AABP; microfibrillar protein;  
 KW abdominal aortic aneurysm disease; treatment; detect; tolerance;  
 KW calcium binding motif.  
 XX  
 OS Unidentified.  
 OS US6048704-A.  
 PN  
 PD 11-APR-2000.  
 XX  
 PF 07-MAR-1997; 97US-0812586.  
 XX  
 PR 07-MAR-1996; 96US-0012976.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Tlison MD;  
 XX  
 DR WPI: 2000-316895/27.  
 XX  
 PS Isolated microfibrillar protein for alleviating abdominal aortic  
 PT aneurysm disease is purified from human aortic tissue and binds  
 PT immunoreactively with immunoglobulin  
 XX  
 PS Example 2; Column 22; 70pp; English.  
 CC The present invention relates to an isolated microfibrillar protein of

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:56:46 : Search time 16.65 Seconds  
(without alignments)  
25.488 Million cell updates/sec

Title: us-09-389-000-2-copy\_163\_169

Perfect score: 1 PYVKGK 7

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size: 0

Total number of hits satisfying chosen parameters: 90768

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database:

A.Geneseq\_0601:\*

1: /cgnl\_9/gcgdata/geneseq/geneseq/AA1980.DAT:\*

2: /cgnl\_9/gcgdata/geneseq/geneseq/AA1981.DAT:\*

3: /cgnl\_9/gcgdata/geneseq/geneseq/AA1982.DAT:\*

4: /cgnl\_9/gcgdata/geneseq/geneseq/AA1983.DAT:\*

5: /cgnl\_9/gcgdata/geneseq/geneseq/AA1984.DAT:\*

6: /cgnl\_9/gcgdata/geneseq/geneseq/AA1985.DAT:\*

7: /cgnl\_9/gcgdata/geneseq/geneseq/AA1986.DAT:\*

8: /cgnl\_9/gcgdata/geneseq/geneseq/AA1987.DAT:\*

9: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT:\*

10: /cgnl\_9/gcgdata/geneseq/geneseq/AA1989.DAT:\*

11: /cgnl\_9/gcgdata/geneseq/geneseq/AA1990.DAT:\*

12: /cgnl\_9/gcgdata/geneseq/geneseq/AA1991.DAT:\*

13: /cgnl\_9/gcgdata/geneseq/geneseq/AA1992.DAT:\*

14: /cgnl\_9/gcgdata/geneseq/geneseq/AA1993.DAT:\*

15: /cgnl\_9/gcgdata/geneseq/geneseq/AA1994.DAT:\*

16: /cgnl\_9/gcgdata/geneseq/geneseq/AA1995.DAT:\*

17: /cgnl\_9/gcgdata/geneseq/geneseq/AA1996.DAT:\*

18: /cgnl\_9/gcgdata/geneseq/geneseq/AA1997.DAT:\*

19: /cgnl\_9/gcgdata/geneseq/geneseq/AA1998.DAT:\*

20: /cgnl\_9/gcgdata/geneseq/geneseq/AA1999.DAT:\*

21: /cgnl\_9/gcgdata/geneseq/geneseq/AA2000.DAT:\*

22: /cgnl\_9/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	57.1	4	21	AA07299
2	4	57.1	5	19	AA061921
3	4	57.1	5	21	AA03688
4	4	57.1	6	18	AA018608
5	4	57.1	6	19	AA061922
6	4	57.1	7	19	AA061923
7	4	57.1	8	21	AA07298
8	4	57.1	9	5	AA040533
9	4	57.1	9	5	AA040534
10	4	57.1	9	13	AA023129
11	4	57.1	9	16	AA080484

12	4	57.1	9	19	AA061925	PPI binding transp
13	4	57.1	9	19	AA061948	PPI binding transp
14	4	57.1	10	14	AA031033	Human IFN-c N-term
15	4	57.1	10	16	AA068941	Human gelsolin ami
16	4	57.1	10	17	AA092146	Porphyromonas ging
17	4	57.1	10	18	AA018606	Aged band 3 peptid
18	4	57.1	10	19	AA061929	PPI-binding peptid
19	4	57.1	10	19	AA061914	Polyporphosinosis
20	4	57.1	10	19	AA052478	Cyclic peptide of
21	4	57.1	10	20	AA025487	Insulin-like growt
22	4	57.1	10	20	AA009668	Antimicrobial cycl
23	4	57.1	10	20	AA099715	Human prostate spe
24	4	57.1	10	21	AA003163	Plasmod pbs240h-g
25	4	57.1	10	21	AA003164	pbs240h-gamma-1FN
26	4	42.9	4	14	AA032984	ICE inhibitor. Sy
27	4	42.9	4	15	AA051048	Sequence of affini
28	3	42.9	4	17	AA077485	Terrapeptide usefu
29	3	42.9	4	17	AA098052	Peptide for use in
30	3	42.9	4	19	AA061920	PPI binding transp
31	3	42.9	4	19	AA056805	Enzyme inhibitor p
32	3	42.9	4	19	AA052403	Beta-turn region u
33	3	42.9	4	20	AA009600	Antimicrobial cycl
34	3	42.9	4	21	AA067506	Caspase 1 peptide
35	3	42.9	4	22	AA030775	Peptide which is u
36	3	42.9	5	13	AA024117	Cyclic pentapeptid
37	3	42.9	5	15	AA069434	Antiinflammatory p
38	3	42.9	5	15	AA062194	CENP-B protein ami
39	3	42.9	5	16	AA071718	Potential cross-li
40	3	42.9	5	16	AA066918	Antagonist peptide
41	3	42.9	5	16	AA066919	Antagonist peptide
42	3	42.9	5	17	AA001950	C140 receptor anta
43	3	42.9	5	17	AA001951	S. solifaciens tr
44	3	42.9	5	17	AA090624	Peptidomimetic inh
45	3	42.9	5	19	AA013860	

#### ALIGNMENTS

RESULT 1	
ID	AA07299 standard; peptide; 4 AA.
AC	AA07299;
XX	
DE	17-JAN-2001 (first entry)
XX	
DE	Synthetic peptide produced by enzymatic amidation.
XX	
KW	Peptide amidation; C-terminal alpha-carboxamide; clostripain;
KW	amidative cleavage; clostripidopeptidase B.
XX	
OS	Synthetic.
XX	
FT	Key
FT	Modified-site 4 Location/Qualifiers
FT	note="C-terminal amide"
XX	
PN	WO200028067-A1.
XX	
PD	18-MAY-2000.
XX	
PF	05-NOV-1999; 99WO-US26060.
XX	
PR	06-NOV-1998; 98US-0107311.
PR	16-DEC-1998; 98US-0212663.
XX	
PA	(BION-) BIONEERASKA INC.
XX	
PI	Dormady D, Scout JS, Strydom DJ, Holmquist B, Wagner FW;
XX	
DR	WPI, 2000-376575/32.
XX	

Search completed: September 15, 2001, 12:50:13  
Job time: 273 sec

Query Match 14.0%; Score 7; DB 14; Length 331;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 YCEQLR 20  
DB 127 YCEQLR 133

RESULT 13  
Q9HV76 PRELIMINARY; PRT; 390 AA.  
AC Q9HV76;  
DT 01-MAR-2001 (TRENBLrel. 16, Created)  
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
DE 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
DE PROBABLE AMINOTRANSFERASE.  
GN PA4722.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sater M.H., Hancock R.B.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen."  
RL Nature 406:959-964(2000).  
DR EMBL; AE004885; AAG08108.1; -.  
DR InterPro; IPR001176; -.  
DR Pfam; PF00155; aminotran\_1; 1.  
DR PROSITE; PS00753; ACCSYNPHASE.  
DR PRSIT; PS00105; AA-TRANSFER\_CLASS\_1; UNKNOWN\_1.  
KW Transferase; Aminotransferase.  
SQ SEQUENCE 390 AA; 42567 MW; CB6A8E94B5C674C3 CRC64;

Query Match 14.0%; Score 7; DB 2; Length 390;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 DAASVLE 38  
DB 215 DAASVLE 221

RESULT 14  
Q9LYR0 PRELIMINARY; PRT; 443 AA.  
AC Q9LYR0;  
DT 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
DE HYPOTHETICAL 50.8 KDA PROTEIN.  
GN T22N19\_110.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,  
RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL163572; CAB87153.1; -.  
DR InterPro; IPR000048; -.  
DR Pfam; PF00612; IQ; 2.  
DR SMART; SM00015; IQ; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 443 AA; 50820 MW; EA3C8B85E2041D83 CRC64;

Query Match 14.0%; Score 7; DB 10; Length 443;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERIK 13  
DB 227 LRRERIK 233

RESULT 15  
O86057 PRELIMINARY; PRT; 500 AA.  
AC O86057;  
DT 01-NOV-1998 (TRENBLrel. 08, Created)  
DT 01-MAR-2001 (TRENBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
DE TRANSCRIPTIONAL ACTIVATOR NTRC.  
GN NTRC.  
OS Herbaspirillum seropedicae.  
OC Bacteria; Proteobacteria; beta subdivision; Oxalobacter group;  
OC Herbaspirillum.  
OX NCBI\_TaxID=964;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Z78;  
RA Steffens M.B.R., Pedrosa F.O., Souza E.M., Persuhn D.C.,  
RA Teixeira K.R.S., Rigo L.U.;  
RT "Sequence and structural organization of the glnAntrBC operon from  
RT Herbaspirillum seropedicae strain Z78."  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER  
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS  
CC -!- SIMILARITY: CONTAINS A SIGMA-34 FACTOR INTERACTION ATP-BINDING  
CC DOMAIN.  
DR EMBL; AF082873; AAC32391.2; -.  
DR HSSP; P41789; INTR.  
DR InterPro; IPR001789; -.  
DR InterPro; IPR002078; -.  
DR InterPro; IPR002197; -.  
DR InterPro; IPR003593; -.  
DR Pfam; PF00072; response\_reg; 1.  
DR Pfam; PF00158; sigma54; 1.  
DR PROSITE; PS00675; SIGMA54\_INTERACT\_1; 1.  
DR PROSITE; PS00676; SIGMA54\_INTERACT\_2; 1.  
DR PROSITE; PS00688; SIGMA54\_INTERACT\_3; 1.  
DR PROSITE; PS50045; SIGMA54\_INTERACT\_4; 1.  
DR SMART; SM00382; AAA; 1.  
DR SMART; SM00448; REC; 1.  
KW ATP-binding; DNA-binding; Phosphorylation; Sensory transduction;  
KW transcription regulation.  
SQ SEQUENCE 500 AA; 54554 MW; 1553B1D8E11816EA CRC64;

Query Match 14.0%; Score 7; DB 2; Length 500;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 GRKNDA 34  
DB 469 GRKNDA 475



ID Q9L8Y3 PRELIMINARY; PRT; 268 AA.  
 AC Q9L8Y3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE VANYB (D-D-CARBOXYPEPTIDASE).  
 GN VANYB OR VANYB2.  
 OS Enterococcus faecalis (Streptococcus faecalis), and  
 OS Enterococcus faecium (Streptococcus faecium).  
 OG Plasmid pIP834.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;  
 OC Enterococcus.  
 OX NCBI\_TaxID=1351, 1352;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.faecalis; STRAIN=BM4382; PLASMID=PIP834; TRANSPOSON=TN1549;  
 RA Garnier F., Taoutit S., Glaser P., Courvalin P., Gallmand M.;  
 RT "Characterization of transposon Tn1549 conferring VanB-type resistance  
 RT in Enterococcus sp.",  
 RL Microbiology 0:0-0(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.faecium; STRAIN=TSGH1;  
 RA Lu J.-J., Perng C.-L., Ho M.-F., Lu C.-L.;  
 RT "Whole sequence of vanB2 gene clusters of Enterococcus faecium  
 RT TSGH1.",  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF192329; AAF72360.1; -;  
 DR EMBL; AF310956; AAG34688.1; -;  
 KW Plasmid; Carboxypeptidase.  
 SQ SEQUENCE 268 AA; 30424 MW; 53FFBC8BDE3B30B1 CRC64;  
  
 Query Match 14.0%; Score 7; DB 2; Length 268;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 34 ASVLEAT 40  
 DB 46 ASVLEAT 52  
  
 RESULT 10  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxID=160;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NICHOLS;  
 RA MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwin E.K., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,  
 RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 RT spirochete.",  
 RL Science 281:375-388(1998).  
 DR EMBL; AE001229; AAC65525.1; -;  
 DR HSSP; P06616; IEGA.  
 DR TIGR; TP0541; -;  
 SQ SEQUENCE 319 AA; 35742 MW; 3DCED9411D1F261A CRC64;

Query Match 14.0%; Score 7; DB 2; Length 319;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 19 LRTLLPY 25  
 DB 287 LRTLLPY 293  
  
 RESULT 11  
 OS Aeropyrum pernix.  
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococaceae;  
 OC Aeropyrum.  
 OX NCBI\_TaxID=56636;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K1;  
 RX MEDLINE=99310339; PubMed=10382966;  
 RA Kavarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Akai A., Kosugi H.,  
 RA Hosoyama A., Fukui Y., Nagai Y., Nishijima K., Nakazawa H.,  
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,  
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 RT crenarchaeon, Aeropyrum pernix K1.",  
 RL DNA Res. 6:83-101(1999).  
 DR EMBL; AP000064; BAA81440.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 326 AA; 35093 MW; 909665F6372DD04C CRC64;  
  
 Query Match 14.0%; Score 7; DB 1; Length 326;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 33 AASVLEA 39  
 DB 292 AASVLEA 298  
  
 RESULT 12  
 OS Molluscum contagiosum virus subtype 1 (MCV1).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Molluscipoxvirus.  
 OX NCBI\_TaxID=10280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,  
 RA Lopez-Esteban J.L., Esteban M., Martin-Gallardo A.;  
 RL Virus Genes 0:0-0(0).  
 DR EMBL; U86899; AAB57946.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 331  
 SQ SEQUENCE 331 AA; 37009 MW; E776B399590B997E CRC64;

```

RP SEQUENCE FROM N.A.
RA Simons G., Groenendijk J., Wijbrandi J., Reijmans M., Groenen J.,
RA Diergaarde P., Van der Lee T., Bleeker M., Onstenk J., De Both M.,
RA Haring M., Mes J., Cornelissen B., Zabeau M., Vos P.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF118127; AAD27815.1; -
DR InterPro; IPR000767; -
DR InterPro; IPR001593; -
DR InterPro; IPR001611; -
DR InterPro; IPR002182; -
DR Pfam; PF00560; LRR; 3; -
DR Pfam; PF00931; NB-ARC; 1;
DR PRINTS; PR00364; DISPASERSIST.
DR PRODOM; PD003035; -; 1.
SQ SEQUENCE 1266 AA; 144826 MW; 8392FFB7F7FD5 CRC64;

Query Match 16.0%; Score 8; DB 10; Length 1266;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 EQLRTLLP 24
DB 547 EQLRTLLP 554
|||||||

RESULT 6
Q9EPS1 PRELIMINARY; PRT; 108 AA.
AC Q9EPS1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PHOSPHODIESTERASE 1A (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Prime G.R., Sutor B.;
RA STRAIN-WISTAR; TISSUE-BRAIN;
RT "Phosphodiesterase 1A (PDE1A) in rat brain.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327836; AAG48734.1; -
FT NON_TER 108
SQ SEQUENCE 108 AA; 12437 MW; 5FD104D5B0CF9D8F CRC64;

Query Match 14.0%; Score 7; DB 11; Length 108;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39
DB 59 AASVLEA 65
|||||||

RESULT 7
O07252 PRELIMINARY; PRT; 200 AA.
AC O07252;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 21.4 KDA PROTEIN.
GN RV0328 OR MTCY63.33.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
CC -!- SIMILARITY: TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL REGULATORS.
DR EMBL; Z96800; CAB09609.1; -
DR Tuberculin; RV0328; -
DR InterPro; IPR001647; -
DR Pfam; PF00440; tetr; 1.
DR PRINTS; PR00455; HTHPTETR.
KW DNA-binding; Hypothetical protein; Transcription regulation.
SQ SEQUENCE 200 AA; 21411 MW; 81ADC84B54B5EC6 CRC64;

Query Match 14.0%; Score 7; DB 2; Length 200;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VLEATVD 42
DB 88 VLEATVD 94
|||||||

RESULT 8
Q9MOV5 PRELIMINARY; PRT; 250 AA.
AC Q9MOV5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 27.6 KDA PROTEIN.
GN AT4G05400.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161503; CAB81082.1; -
KW Hypothetical protein.
SQ SEQUENCE 250 AA; 27627 MW; 7AAA04FBF0A6F707 CRC64;

Query Match 14.0%; Score 7; DB 10; Length 250;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSSEKLR 8
DB 244 SSSEKLR 250
|||||||

RESULT 9
Q9L8Y3

```

GN DRC0037.  
 OS Deinococcus radiodurans.  
 OG Plasmid Cpl.  
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.  
 CX NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RL;  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamathevan J.J., Lam P., McDonald L., Otterback T., Zalewski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 RT radiodurans RL"; 1577(1999).  
 RL Science 286:1571-1577(1999).  
 DR EMBL; AEO01827; AAF12674.1; -.  
 DR TIGR; DRC0037; -.  
 KW Plasmid.  
 SQ SEQUENCE 703 AA; 77447 MW; 6ED24979D0ECD646 CRC64;  
  
 QY 33 AASVLEAT 40  
 | | | | | | | |  
 DB 662 AASVLEAT 669  
  
 RESULT 3  
 Q24015  
 ID Q24015 PRELIMINARY; PRT; 1220 AA.  
 AC Q24015;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE RESISTANCE COMPLEX PROTEIN I2C-1.  
 GN I2C-1.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
 OC Solanales; Solanaceae; Solanum.  
 CX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97290204; PubMed=9144960;  
 RA Ori N., Eshed Y., Paran I., Presting G., Aviv D., Tanksley S.,  
 RA Zamir D., Fluhr R.;  
 RT "The I2C family from the wilt disease resistance locus I2 belongs to  
 RT the nucleotide binding, leucine-rich repeat superfamily of plant  
 RT resistance genes";  
 RL Plant Cell 9:521-532(1997).  
 RL EMBL; AF004878; AAB63274.1; -.  
 DR Mendel; 24472; Lyses; 3172; 24472.  
 DR InterPro; IPR000767; -.  
 DR InterPro; IPR001611; -.  
 DR InterPro; IPR002182; -.  
 DR InterPro; IPR003592; -.  
 DR Pfam; PF00550; LRR; 4.  
 DR Pfam; PF00931; NB-ARC; 1.  
 DR PRINTS; PR00364; DISEASERISIST.  
 DR SMART; SM00370; LRR; 1.  
 SQ SEQUENCE 1220 AA; 139386 MW; B5997280C8C5EB20 CRC64;  
  
 Query Match 16.0%; Score 8; DB 10; Length 1220;  
 Best Local Similarity 100.0%; Pred. No. 8.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:50:11 ; Search time 69.13 Seconds  
(without alignments)  
95.693 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_189

Perfect score: 50

Sequence: 1 HSSKELRRRIKYCEQLR.....NDAASVLEATVDYVKYIREK 50

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 425036 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL16:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_unclassified:\*
- 13: sp\_vertebrate:\*
- 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	88.0	425	4 Q9NX45	Q9NX45 homo sapien
2	8	16.0	703	2 Q9RZF3	Q9RZF3 deinococcus
3	8	16.0	1220	10 Q24015	Q24015 lycopersico
4	8	16.0	1240	10 Q24016	Q24016 lycopersico
5	8	16.0	1266	10 Q9XET3	Q9XET3 lycopersico
6	7	14.0	108	11 Q9EPS1	Q9EPS1 rattus norv
7	7	14.0	200	2 O07252	O07252 mycobacteri
8	7	14.0	250	10 Q9M0V5	Q9M0V5 arabidopsis
9	7	14.0	268	2 Q9L8Y3	Q9L8Y3 enterococcu
10	7	14.0	319	2 Q83552	Q83552 treponema p
11	7	14.0	326	1 Q9Y960	Q9Y960 aeropyrum p
12	7	14.0	331	14 Q11316	Q11316 molluscum c
13	7	14.0	390	2 Q9HV76	Q9HV76 pseudomonas
14	7	14.0	443	10 Q9LYR0	Q9LYR0 arabidopsis
15	7	14.0	500	2 Q86057	Q86057 herbaspiril
16	7	14.0	502	2 Q9ZB31	Q9ZB31 bruceella ab
17	7	14.0	514	6 Q28063	Q28063 bos taurus
18	7	14.0	519	4 Q9UEFX3	Q9UEFX3 homo sapien
19	7	14.0	542	11 Q9EPR9	Q9EPR9 rattus norv

20	7	14.0	586	2 Q9ZDF1	Q9ZDF1 rickettsia
21	7	14.0	746	10 Q9FNF6	Q9FNF6 arabidopsis
22	7	14.0	748	14 Q88200	Q88200 molluscum c
23	7	14.0	791	5 Q45358	Q45358 caenorhabdi
24	7	14.0	974	2 Q84831	Q84831 chlamydia t
25	7	14.0	1080	11 Q9WVP5	Q9WVP5 mus musculu
26	7	14.0	1080	11 Q9EQU1	Q9EQU1 mus musculu
27	7	14.0	1223	5 Q9VRC9	Q9VRC9 drosophila
28	7	14.0	1260	5 Q9GRG5	Q9GRG5 trypanosoma
29	7	14.0	1270	13 Q12982	Q12982 xenopus lae
30	7	14.0	1341	3 Q08281	Q08281 saccharomyc
31	7	14.0	1342	3 Q92271	Q92271 saccharomyc
32	7	14.0	1833	5 Q9VM67	Q9VM67 drosophila
33	6	12.0	39	4 Q9UBN9	Q9UBN9 homo sapien
34	6	12.0	58	9 Q38502	Q38502 bacterioph
35	6	12.0	64	2 Q9EY25	Q9EY25 photobacter
36	6	12.0	81	2 Q52484	Q52484 burkholderi
37	6	12.0	87	2 Q9KV11	Q9KV11 vibrio chol
38	6	12.0	88	14 Q9EAM8	Q9EAM8 human immun
39	6	12.0	90	1 Q26558	Q26558 methanobact
40	6	12.0	94	1 Q9HH63	Q9HH63 methanobact
41	6	12.0	94	9 Q80193	Q80193 methanobact
42	6	12.0	106	2 Q9KIT8	Q9KIT8 bacillus me
43	6	12.0	109	2 Q9KP60	Q9KP60 vibrio chol
44	6	12.0	111	2 Q06227	Q06227 borrelia bu
45	6	12.0	111	2 Q06230	Q06230 borrelia bu

#### ALIGNMENTS

RESULT 1

Q9NX45 ID Q9NX45 PRELIMINARY; PRT; 425 AA.  
AC Q9NX45;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
DE CDNA FLJ20449 FIS, CLONE KAT05575.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,  
RA Suzuki Y., Ohayashi M., Nishi T., Shibahara T., Tanaka T.,  
RA Nakamura Y., Isono T., Sugano S.;  
RT "NEDO human cDNA sequencing project."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK000456; BAA91175.1; -.  
DR InterPro; IPR001092; -.  
DR SMART; SM00353; HLH; 1.  
SQ SEQUENCE 425 AA; 46922 MW; 777E57231F9D6BDD CRC64;

Query Match 88.0%; Score 44; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 3e-37;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRERIKYCCCLRLPLPVKGRKNDAAASVLEATVDYVKYIREK 50

Db 212 LRRERIKYCCCLRLPLPVKGRKNDAAASVLEATVDYVKYIREK 255

RESULT 2

Q9RZF3 ID Q9RZF3 PRELIMINARY; PRT; 703 AA.  
AC Q9RZF3;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
DE NODULATION PROTEIN-RELATED PROTEIN.

Search completed: September 15, 2001, 12:50:41  
Job time: 286 sec

---

Query Match 12.0%; Score 6; DB 1; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERI 12  
 |||||  
 DB 14 LRRERI 19

## RESULT 14

YB4C\_SCHPO STANDARD; PRT; 93 AA.  
 AC O14358;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HYPOTHETICAL 10.7 KDA PROTEIN C30D10.12C IN CHROMOSOME II.  
 GN SPBC30D10.12C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Wood V., Rajandream M.A., Barrell B.G., Lauber J., Hilbert H.,  
 RA Dusterhoeft A.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: TO YEAST YGR215W.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z97992; CAB10807.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 93 AA; 10698 MW; 6ECC74F169DA747B CRC64;

Query Match 12.0%; Score 6; DB 1; Length 93;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 QLRTLL 23  
 |||||  
 DB 55 QLRTLL 60

## RESULT 15

ASFP\_BOVIN STANDARD; PRT; 134 AA.  
 AC P29392;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE ACIDIC SEMINAL FLUID PROTEIN PRECURSOR (ASFP).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seminal vesicle;  
 RX MEDLINE=92181448; PubMed=1543494;  
 RA Wempe F., Einspanier R., Scheit K.H.;  
 RT "Characterization by cDNA cloning of the mRNA of a new growth factor  
 RT from bovine seminal plasma: acidic seminal fluid protein.";

RL Biochem. Biophys. Res. Commun. 183:232-237(1992).  
 RN [2]  
 RP SEQUENCE OF 21-43.  
 RC TISSUE=Seminal vesicle;  
 RX MEDLINE=91378963; PubMed=1898381;  
 RA Einspanier R., Einspanier A., Wempe F., Scheit K.H.;  
 RT "Characterization of a new bioactive protein from bovine seminal  
 RT fluid.";  
 RL Biochem. Biophys. Res. Commun. 179:1006-1010(1991).  
 RN [3]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=94237283; PubMed=8181566;  
 RA Einspanier R., Krause I., Calvete J.J., Toepfer-Petersen E.,  
 RA Klostermeyer H., Karg H.;  
 RT "Bovine seminal plasma ASFP: localization of disulfide bridges and  
 RT detection of three different isoelectric forms.";  
 RL FEBS Lett. 344:61-64(1994).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=97475216; PubMed=9334740;  
 RA Romero A., Romao M.J., Varela P.F., Koellin I., Dias J.M.,  
 RA Carvalho A.L., Sanz L., Toepfer-Petersen E., Calvete J.J.;  
 RT "The crystal structures of two spermadhesins reveal the CUB domain  
 RT fold.";  
 RL Nat. Struct. Biol. 4:783-788(1997).  
 CC -!- FUNCTION: STIMULATES CELL DIVISION AND PROGESTERONE SECRETION  
 CC OF BOVINE GRANULOSA CELLS IN VITRO IN A POTENT AND DOSE DEPENDENT  
 CC MANNER. THIS PROTEIN APPEARS TO BE A POTENT GROWTH FACTOR WITH  
 CC EFFECTS ON OVARIAN GRANULOSA CELLS.  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC -!- TISSUE SPECIFICITY: SEMINAL VESICLE TISSUE, AMPULLA AND WEAKLY  
 CC IN TISSUE OF EPIDIDYMIS.  
 CC -!- SIMILARITY: CONTAINS 1 CUB DOMAIN.  
 CC -!- SIMILARITY: BELONGS TO THE SPERMADHESIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M84603; AAA30745.1; -.  
 DR PIR; PH0213; PH0213.  
 DR PIR; JQ1403; JQ1403.  
 DR PDB; 1SFP; 24-JUN-98.  
 DR InterPro; IPR000124; -.  
 DR InterPro; IPR000859; -.  
 DR Pfam; PF00431; CUB; 1.  
 DR PROSITE; PS00985; SPERMADHESIN\_1; 1.  
 DR PROSITE; PS00986; SPERMADHESIN\_2; 1.  
 DR PROSITE; PS01180; CUB; 1.  
 KW Growth factor; Signal; 3D-structure.  
 FT SIGNAL 1 20  
 FT CHAIN 21 134 ACIDIC SEMINAL FLUID PROTEIN.  
 FT DOMAIN 30 51  
 FT DISULFID 30 51  
 FT DISULFID 74 95  
 FT CONFLICT 43 43 T -> H (IN REF. 2).  
 SQ SEQUENCE 134 AA; 15036 MW; 339BCFF8637D64C0 CRC64;

Query Match 12.0%; Score 6; DB 1; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 VKYIRE 49  
 |||||  
 DB 111 VKYIRE 116

RT "The proteinase yscH-inhibitor, IA3, gene. Studies of cytoplasmic  
 RL proteinase inhibitor deficiency on yeast physiology.",  
 RN FEBS Lett. 283:78-84(1991).  
 RP (2)  
 RC SEQUENCE FROM N.A.  
 RA STRAIN-S288C / AB972;  
 RL Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.,  
 RA Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RP (3)  
 RN SEQUENCE.  
 RA Biedermann K., Montali U., Martin B., Svendsen I., Ottesen M.,  
 RT "The amino acid sequence of proteinase A inhibitor 3 from baker's  
 RL yeast".  
 CC Carlsberg Res. Commun. 45:225-235(1980).  
 CC -!- FUNCTION: INHIBITOR FOR YEAST PROTEASE A (YSCA).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; X60050; CAA42650.1; -;  
 DR EMBL; Z49808; CAA89907.1; -;  
 DR PIR; A01334; IABY3.  
 DR PIR; S16692; S16692.  
 DR SGD; S0004786; PAI3.  
 KW Protease inhibitor; Acetylation.  
 FT MOD\_RES 1 1 ACETYLATION.  
 FT DOMAIN 2 23 NEARLY ALL OF THE INHIBITORY ACTIVITY IS  
 FT PRESENT IN THIS REGION.  
 FT SEQUENCE 68 AA; 7707 MW; 8295BF7AF3E4A01A CRC64;  
 SQ  
 Query Match 12.0%; Score 6; DB 1; Length 68;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 SSKEKL 7 PRT; 79 AA.  
 DB 14 SSKEKL 19  
 RESULT 12  
 RS18\_UREPA ID\_UREPA STANDARD; PRT; 79 AA.  
 AC Q9PPT8; DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE 30S RIBOSOMAL PROTEIN S18.  
 GN RPSR OR RPS18 OR U0552.  
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 CC Mycoplasmataceae; Ureaplasma.  
 CC NCBI\_TaxID=134821;  
 CC (1)  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-SEROVAR 3;  
 RX MEDLINE=20500219; PubMed=11048724;  
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
 RA Cassell G.H.;  
 RT "The complete sequence of the mucosal pathogen Ureaplasma  
 RL urealyticum".  
 RL Nature 407:757-762(2000).  
 CC -!- FUNCTION: THIS PROTEIN HAS BEEN IMPLICATED IN AMINOACYL-TRANSFER  
 CC RNA BINDING. IT APPEARS TO BE SITUATED AT THE DECODING SITE OF  
 CC MESSENGER RNA (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AE002154; AAF30965.1; -;  
 DR InterPro; IPR001648; -;  
 DR Pfam; PF01084; Ribosomal\_S18; 1.  
 DR PRINTS; PR00974; RIBOSOMALS18  
 DR PROSITE; PS00057; RIBOSOMAL\_S18; FALSE\_NEG.  
 KW Ribosomal protein; RNA-binding.  
 SQ SEQUENCE 79 AA; 9136 MW; E4697FD3F03F5AC4 CRC64;  
 Query Match 12.0%; Score 6; DB 1; Length 79;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 22 LLPYVK 27  
 DB 73 LLPYVK 78  
 RESULT 13  
 HFQ\_HAEIN ID\_HFQ\_HAEIN STANDARD; PRT; 90 AA.  
 AC P44437; DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE HOST FACTOR-I PROTEIN (HF-I).  
 GN HFO OR HI0411.  
 OS Haemophilus influenzae.  
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 CC Haemophilus.  
 CC NCBI\_TaxID=727;  
 CC (1)  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 RL influenzae Rd".  
 RL Science 269:496-512(1995).  
 CC -!- FUNCTION: REQUIRED FOR PHASE Q BETA RNA-DIRECTED SYNTHESIS OF  
 CC COMPLEMENTARY MINUS-STRAND RNA (BY SIMILARITY).  
 CC -!- SUBUNIT: HOMOHXAMER (BY SIMILARITY).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U32724; AAC22070.1; -;  
 DR TIGR; HI0411; -;  
 KW RNA-binding.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 SQ SEQUENCE 90 AA; 10118 MW; 1C91D4ECF1F7F029 CRC64;

SUBFAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; X64603; CAA45887.1; -;  
DR PIR; A44337; A44337.  
DR HSSP; P17119; 3KAR.  
DR InterPro; IPR001752; -;  
DR Pfam; PF00225; kinesin; 1.  
DR PRINTS; PR00380; KINESINHEAVY.  
DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
DR PROSITE; PS00067; KINESIN\_MOTOR\_DOMAIN2; 1.  
KW Motor protein; Microtubules; ATP-binding; Coiled coil.  
FT DOMAIN 175 425 COILED COIL (POTENTIAL).  
FT DOMAIN 426 770 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).  
FT NP\_BIND 514 521 ATP (BY SIMILARITY).  
SQ SEQUENCE 770 AA; 85800 MW; EFD0F0FF39B5C7EB CRC64;

Query Match 14.0%; Score 7; DB 1; Length 770;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KEKLRE 10  
| | | | | | |  
Db 398 KEKLRE 404

RESULT 9  
ID YMC9\_YEAST STANDARD; PRT; 838 AA.  
AC Q03714;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE HYPOTHETICAL 96.7 KDA PROTEIN IN NDC1-TSAL INTERGENIC REGION.  
GN YML029W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; Z46659; CAA86626.1; -;  
DR SGD; S0004491; YML029W.  
DR InterPro; IPR000626; -;  
DR PROSITE; PS00053; UBIQUITIN\_2; UNKNOWN\_1.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 224 244 POTENTIAL.  
FT DOMAIN 259 318 UBIQUITIN-LIKE.  
FT TRANSMEM 532 552 POTENTIAL.  
FT TRANSMEM 553 573 POTENTIAL.  
FT TRANSMEM 764 784 POTENTIAL.

SQ SEQUENCE 838 AA; 96653 MW; 9B93ECA6C5421FD6 CRC64;

Query Match 14.0%; Score 7; DB 1; Length 838;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 TLLPYVK 27  
| | | | | | |  
Db 778 TLLPYVK 784

RESULT 10  
VG1A\_BPP2A STANDARD; PRT; 59 AA.  
ID VG1A\_BPP2A  
AC P06947;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 01-JAN-1988 (Rel. 06, Last annotation update)  
DE EARLY PROTEIN GPIA.  
GN 1A.  
OS Bacteriophage PZA.  
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Podoviridae.  
OX NCBI\_TaxID=10757;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=8605991; PubMed=3934048;  
RA Paces V., Vitek C., Urbanek P., Hostomsky Z.;  
RT "Nucleotide sequence of the major early region of Bacillus subtilis phage PZA, a close relative of phi 29.";  
RL Gene 38:45-56(1985).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; M11813; AAA88481.1; -;  
DR PIR; A24528; ERBP1A.  
KW Early protein.  
SQ SEQUENCE 59 AA; 6865 MW; FCC525137B72D831 CRC64;

Query Match 12.0%; Score 6; DB 1; Length 59;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 DAASVL 37  
| | | | | | |  
Db 28 DAASVL 33

RESULT 11  
IPA3\_YEAST STANDARD; PRT; 68 AA.  
ID IPA3\_YEAST  
AC P01094;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE PROTEASE A INHIBITOR 3 (PROTEINASE INHIBITOR I(A)3).  
GN PA13 OR YMR174C OR YW010.04C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=S288C;  
RC MEDLINE=91243884; PubMed=2037077;  
RA Schu P., Wolf D.H.;



GN POTA OR MG042.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.:  
 RT "The minimal gene complement of Mycoplasma genitalium";  
 RL Science 270:397-403(1995).  
 CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
 CC FOR SPERMIDINE/PUTRESCINE. PROBABLY RESPONSIBLE FOR ENERGY  
 CC COUPLING TO THE TRANSPORT SYSTEM (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED (POTENTIAL).  
 CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; U39684; AAC71258.1; -.  
 DR HSSP; P13569; INBD.  
 DR TIGR; MG042; -.  
 DR InterPro: IPR001617; -.  
 DR Pfam: PF00005; ABC\_tran; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 DR Transport; ATP-binding; Membrane.  
 KW NP\_BIND 40 47 ATP (POTENTIAL).  
 FT NP\_SEQUENCE 559 AA; 65192 MW; DB8F04B26A110F6F CRC64;  
 SQ SEQUENCE 559 AA; 65192 MW; DB8F04B26A110F6F CRC64;  
 Query Match 14.0%; Score 7; DB 1; Length 559;  
 Best Local Similarity 100.0%; Pred. No. 8.6;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 KEKLRE 10  
 DB 173 KEKLRE 179  
 RESULT 7  
 CNIA\_MOUSE STANDARD; PRT; 565 AA.  
 ID CNIA\_MOUSE  
 AC Q61481; O35388;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DE 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CALCIUM/CALMODULIN-DEPENDENT 3',5'-CYCLIC NUCLEOTIDE PHOSPHODIESTERASE  
 DE 1A (EC 3.1.4.17) (CAM-PDE 1A) (61 KDA CAM-PDE).  
 GN PDE1A.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (PDE1A2).  
 RC STRAIN=BALB/C; TISSUE=Brain;  
 RC Ian C., Sonnenburg W.K., Zhao A.2., Kwak K.S., Beavo J.A.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RP SEQUENCE OF 1-262 FROM N.A. (PDE1A1).

RC TISSUE=Heart;  
 RA Sonnenburg W.K., Rybalkin S.D., Bornfeldt K.E., Kwak K.S.,  
 RA Rybalkina I., Beavo J.A.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: HAS A HIGHER AFFINITY FOR GMP THAN FOR CAMP.  
 CC -!- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =  
 CC GUANOSINE 5'-PHOSPHATE.  
 CC -!- ENZYME REGULATION: TYPE I PDE ARE ACTIVATED BY THE BINDING OF  
 CC CALMODULIN IN THE PRESENCE OF CA(2+).  
 CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -!- ALTERNATIVE PRODUCTS: THERE ARE DIFFERENT ISOFORMS PRODUCED BY  
 CC ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ISOFORM  
 CC PDE1A2.  
 CC -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; U56649; AAB03319.1; -.  
 DR EMBL; AF023529; AAB81952.1; -.  
 DR MGD; MGI:1201792; Pdela.  
 DR InterPro: IPR002073; -.  
 DR Pfam: PF00233; PDEase; 1.  
 DR PRINTS; PR00387; PDIESTERASE1.  
 DR PROSITE; PS00126; PDEASE\_1; 1.  
 DR HYDROLASE; CAMP; CGMP; Multigene family; Alternative splicing;  
 KW Calmodulin-binding.  
 FT DOMAIN 44 64 CALMODULIN-BINDING.  
 FT DOMAIN 213 535 CATALYTIC (BY SIMILARITY).  
 FT VARSPLIC 1 54 MYGSSSTSSSHWIAVRNIMGSTDTDELENATYKYLIG  
 FT EOTERKMWORLKG I -> MDEYVTIRKKHLQRPFR (IN  
 FT ISOFORM PDE1A1).  
 SQ SEQUENCE 565 AA; 64470 MW; 56A0749774967FE6 CRC64;  
 Query Match 14.0%; Score 7; DB 1; Length 565;  
 Best Local Similarity 100.0%; Pred. No. 8.6;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 33 AASVLEA 39  
 DB 79 AASVLEA 85  
 RESULT 8  
 KLPA\_EMENI STANDARD; PRT; 770 AA.  
 ID KLPA\_EMENI  
 AC P28739;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE KINESIN-LIKE PROTEIN KLPA.  
 GN KLPA.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; Emericella.  
 OX NCBI\_TaxID=5072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GB20;  
 RX MEDLINE=93107178; PubMed=8416986;  
 RA O'Connell M.J., Meluh P.B., Rose M.D., Morris N.R.;  
 RT "Suppression of the bimC4 mitotic spindle defect by deletion of klpa,  
 RT a gene encoding a KAR3-related kinesin-like protein in Aspergillus  
 RT nidulans.";  
 RL J. Cell Biol. 120:153-162(1993).  
 RL -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD

```

RX MEDLINE-93107074; PubMed-7678006;
RA Sonnenburg W.K., Seger D., Beavo J.A.;
RT "Molecular cloning of a cDNA encoding the '61-kDa'
RT calmodulin-stimulated cyclic nucleotide phosphodiesterase."
RT Tissue-specific expression of structurally related isoforms."
RL J. Biol. Chem. 268:645-652(1993).
RN [2]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE-91329365; PubMed-1651111;
RA Charbonneau H., Kumar S., Novack J.P., Blumenthal D.K., Griffin P.R.,
RA Shabanowitz J., Hunt D.F., Beavo J.A., Walsh K.A.;
RT "Evidence for domain organization within the 61-kDa
RT calmodulin-dependent cyclic nucleotide phosphodiesterase from bovine
RT brain."
RL Biochemistry 30:7931-7940(1991).
RN [3]
RP SEQUENCE OF 193-426.
RC TISSUE=Brain;
RX MEDLINE-87092242; PubMed-3025833;
RA Charbonneau H., Beier N., Walsh K.A., Beavo J.A.;
RT "Identification of a conserved domain among cyclic nucleotide
RT phosphodiesterases from diverse species."
RL Proc. Natl. Acad. Sci. U.S.A. 83:9308-9312(1986).
CC -!- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC GUANOSINE 5'-PHOSPHATE.
CC -!- ENZYME REGULATION: TYPE I PDE ARE ACTIVATED BY THE BINDING OF
CC CALMODULIN IN THE PRESENCE OF CA(2+).
CC -!- SUBUNIT: HOMODIMER.
CC -!- ALTERNATIVE PRODUCTS: THERE ARE PROBABLY DIFFERENT ISOFORMS
CC PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF
CC ISOFORM PDE1A2.
CC -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M90358; AAA74560.1;
CC PIR; A26650; A26650.
CC PIR; A40282; A40282.
CC PIR; A45334; A45334.
CC InterPro; IPR002073;
CC Pfam; PF00233; PDEase_1.
CC PRINTS; PR00387; PDIESTERASE1.
CC PROSITE; PS00126; PDEase_1.
CC Hydrolase; CAMP; cGMP; Multigene family; Alternative splicing;
CC Calmodulin-binding.
CC INIT_MET 0 0
CC DOMAIN 23 43 CALMODULIN-BINDING.
CC CONFLICT 192 500 CATALYTIC (BY SIMILARITY).
CC CONFLICT 236 500 H -> G (IN REF. 3).
CC CONFLICT 320 320 N -> W (IN REF. 3).
CC SEQUENCE 529 AA; 60712 MW; 24DB74C22A2AE06F CRC64;

Query Match 14.0%; Score 7; DB 1; Length 529;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39
Db 58 AASVLEA 64
|||||||
RESULT 5
CN1A_HUMAN STANDARD; PRT; 534 AA.
ID

```

```

AC P54750;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CALCIUM/CALMODULIN-DEPENDENT 3',5'-CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
DE 1A (EC 3.1.4.17) (CAM-PDE 1A) (61 KDA CAM-PDE) (HCAM-1).
GN PDE1A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96132810; PubMed-8557689;
RA Loughney K., Martins T.J., Harris E.A.S., Sadhu K., Hicks J.B.,
RA Sonnenburg W.K., Beavo J.A., Ferguson K.;
RT "Isolation and characterization of cDNAs corresponding to two human
RT calcium, calmodulin-regulated, 3',5'-cyclic nucleotide
RT phosphodiesterases."
RL J. Biol. Chem. 271:796-806(1996).
CC -!- FUNCTION: HAS A HIGHER AFFINITY FOR CGMP THAN FOR CAMP.
CC -!- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC GUANOSINE 5'-PHOSPHATE.
CC -!- ENZYME REGULATION: TYPE I PDE ARE ACTIVATED BY THE BINDING OF
CC CALMODULIN IN THE PRESENCE OF CA(2+).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- ALTERNATIVE PRODUCTS: THERE ARE PROBABLY DIFFERENT ISOFORMS
CC PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF
CC ISOFORM PDE1A3.
CC -!- TISSUE SPECIFICITY: SEVERAL TISSUES, INCLUDING BRAIN, KIDNEY,
CC TESTIS, AND HEART.
CC -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U40370; AAC50436.1;
CC MIM; 171890;
CC InterPro; IPR002073;
CC Pfam; PF00233; PDEase_1.
CC PRINTS; PR00387; PDIESTERASE1.
CC PROSITE; PS00126; PDEase_1.
CC Hydrolase; CAMP; cGMP; Multigene family; Alternative splicing;
CC Calmodulin-binding.
CC INIT_MET 0 0 BY SIMILARITY.
CC DOMAIN 23 43 CALMODULIN-BINDING (BY SIMILARITY).
CC CONFLICT 192 514 CATALYTIC (BY SIMILARITY).
CC SEQUENCE 534 AA; 61120 MW; 8398FC45160BA720 CRC64;

Query Match 14.0%; Score 7; DB 1; Length 534;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39
Db 58 AASVLEA 64
|||||||
RESULT 6
POTA_MYCGE STANDARD; PRT; 559 AA.
ID
AC P47288;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SPERMIDINE/PUTRESCINE TRANSPORT ATP-BINDING PROTEIN POTA HOMOLOG.
DE

```

DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HYPOTHETICAL PROTEIN PA4543.  
 GN PA4543.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goulet L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964 (2000).  
 RN [2]  
 RP SEQUENCE OF 198-242 FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=93225810; PubMed=8097014;  
 RA Hobbs M., Collie E.S., Free P.D., Livingston S.P., Mattick J.S.;  
 RT "Pils and PilR, a two-component transcriptional regulatory system  
 RT controlling expression of type 4 fimbriae in Pseudomonas  
 RT aeruginosa.";  
 RL Mol. Microbiol. 7:669-682 (1993).  
 CC -!- SIMILARITY: BELONGS TO THE UPF0124 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; AB004868; AAG07931.1; -  
 DR EMBL; L06013; AAB87639.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 242 AA; 26047 MW; AD4470FF676FF51D CRC64;  
 Query Match 14.0%; Score 7; DB 1; Length 242;  
 Best Local Similarity 100.0%; Pred. No. 4.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 36 VLEATVD 42  
 Db 130 VLEATVD 136  
 |||||  
 RESULT 3  
 ID DPPE\_HAEIN STANDARD; PRT; 327 AA.  
 AC P45094;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE DIPEPTIDE TRANSPORT ATP-BINDING PROTEIN DPPE.  
 GN DPPE OR H1184.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McElroy K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 RT influenzae Rd.";  
 RL Science 269:496-512 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20 / ATCC 51907;  
 RX MEDLINE=96134971; PubMed=8550458;  
 RA Preston A., Maskell D., Johnson A., Moxon E.R.;  
 RT "Altered lipopolysaccharide characteristic of the 169 phenotype in  
 RT Haemophilus influenzae results from mutations in a novel gene, isn.";  
 RL J. Bacteriol. 178:396-402 (1996).  
 CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
 CC FOR DIPEPTIDES. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO  
 CC THE TRANSPORT SYSTEM (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED (POTENTIAL).  
 CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; U32798; AAC22837.1; -  
 DR EMBL; U17295; AAA95975.1; -  
 DR TIGR; H11184; -  
 DR InterPro; IPR001617; -  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 KW Peptide transport; Transport; Inner membrane; ATP-binding.  
 FT NP\_BIND 34 61  
 FT SEQUENCE 327 AA; 36917 MW; OBH0BDE197DA9BE CRC64;  
 Query Match 14.0%; Score 7; DB 1; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 5.4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 LRRERIK 13  
 Db 271 LRRERIK 277  
 |||||  
 RESULT 4  
 ID CNIA\_BOVIN STANDARD; PRT; 529 AA.  
 AC P14100;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CALCIUM/CALMODULIN-DEPENDENT 3',5'-CYCLIC NUCLEOTIDE PHOSPHODIESTERASE  
 DE 1A (EC 3.1.4.17) (CAM-PDE 1A) (61 KDA CAM-PDE).  
 GN PDE1A.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:50:40 : Search time 23.18 seconds  
(without alignments)  
73.890 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_189

Perfect score: 50

Sequence: 1 HSSREKLRRRIKVCCEQLR.....NDAASVLEATVDYVKYIREK 50

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	14.0	169	1 Y769_METJA	Q58179 methanococ
2	7	14.0	242	1 Y9E3_PSEAE	P33663 pseudomonas
3	7	14.0	327	1 DPPE_HAEIN	P45094 haemophilus
4	7	14.0	529	1 CNIA_BOVIN	P14100 bos taurus
5	7	14.0	534	1 CNIA_HUMAN	P54750 homo sapien
6	7	14.0	559	1 POTA_MYCSE	P47288 mycoplasma
7	7	14.0	565	1 CNIA_MOUSE	Q61481 mus muscula
8	7	14.0	770	1 KLPA_EMENI	P28739 emericella
9	7	14.0	838	1 YMC9_YEAST	Q03714 saccharomyc
10	6	12.0	59	1 VGIA_BPPZA	P06947 bacterioph
11	6	12.0	68	1 IPA3_YEAST	P01094 saccharomyc
12	6	12.0	79	1 RS18_UREPA	Q99PT8 ureaplasma
13	6	12.0	90	1 HFQ_HAEIN	P44437 haemophilus
14	6	12.0	93	1 YB4C_SCHPO	O14358 schizosacch
15	6	12.0	134	1 ASPP_BOVIN	P29392 bos taurus
16	6	12.0	145	1 RR17_ORYSA	Q9ZST1 oryza sativ
17	6	12.0	145	1 Y004_TREPA	O83050 treponema p
18	6	12.0	150	1 LCXS_HUMAN	Q16873 homo sapien
19	6	12.0	153	1 MAL_CANFA	Q28296 canis fami
20	6	12.0	153	1 MAL_HUMAN	P21145 homo sapien
21	6	12.0	153	1 MAL_MOUSE	O09198 mus musculu
22	6	12.0	153	1 MAL_RAT	Q64349 rattus norv
23	6	12.0	159	1 GREB_BUCAI	P74464 buccinera ap
24	6	12.0	160	1 YMT0_YEAST	Q04210 saccharomyc
25	6	12.0	177	1 PUR6_PYRHO	O58058 pyrococcus
26	6	12.0	181	1 NOHB_ECOLI	P31062 escherichia
27	6	12.0	181	1 TERS_LAMB	P03707 bacterioph
28	6	12.0	189	1 NOHA_ECOLI	P31061 escherichia
29	6	12.0	239	1 RP35_BACTK	P26763 bacillus th
30	6	12.0	242	1 YTXE_BACSU	P39064 bacillus su
31	6	12.0	248	1 DSBG_ECOLI	P77202 escherichia
32	6	12.0	250	1 VGLL_HSV6U	P52508 human herpe
33	6	12.0	250	1 VGLL_HSV6Z	P52526 human herpe

34	6	12.0	258	1 VGLL_GPCMV	O92277 guinea pig
35	6	12.0	268	1 RPNA_YEAST	P38886 saccharomyc
36	6	12.0	268	1 VANY_ENTFA	O47746 enterococcu
37	6	12.0	274	1 OSA4_BORBU	Q04851 borrelia bu
38	6	12.0	274	1 YA99_SCHPO	Q09787 schizosacch
39	6	12.0	289	1 IPYR_BOVIN	P37980 bos taurus
40	6	12.0	289	1 IPYR_HUMAN	Q15181 homo sapien
41	6	12.0	291	1 YXJO_BACSU	P55181 bacillus su
42	6	12.0	292	1 YN19_MYCTU	P71893 mycobacteri
43	6	12.0	318	1 YZ11_AQUAE	O66405 aquifex aeo
44	6	12.0	328	1 CEBB_CHICK	Q05826 gallus gall
45	6	12.0	334	1 DPPE_ECOLI	P37313 escherichia

ALIGNMENTS

RESULT 1					
ID Y769_METJA	STANDARD;	PRT;	169 AA.		
AC Q58179;					
DT 01-NOV-1997 (Rel. 35, Created)					
DT 01-NOV-1997 (Rel. 35, Last sequence update)					
DE 01-NOV-1997 (Rel. 35, Last annotation update)					
DE HYPOTHETICAL PROTEIN MJ0769.					
GN MJ0769.					
OS Methanococcus jannaschii.					
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;					
OC Methanococcus.					
OX NCBI_TaxID=2190;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;					
RX MEDLINE=96337999; PubMed=8688087;					
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,					
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Adams M.D., Reich C.I.,					
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reick J.M., Glodek A.,					
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,					
RA Scott J.L., Geoghegan N.S.M., Weidman J.D., Sadow P.W., Hanna M.C.,					
RA Utterback T.R., Kelley J.M., Peterson J.D., Hurd M.A., Kaine B.P., Borodovsky M.,					
RA Klenk H.-P., Roberts C.M., Smith H.O., Woese C.R., Venter J.C.;					
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus					
RL jannaschii".					
RL Science 273:1058-1073(1996).					
CC -----					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC the European Bioinformatics Institute. There are no restrictions on its					
CC use by non-profit institutions as long as its content is in no way					
CC modified and this statement is not removed. Usage by and for commercial					
CC entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>					
CC or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).					
CC -----					
DR EMBL; U67522; AAB98774.1; ..					
DR TIGR; MJ0769; ..					
KW Hypothetical protein.					
SQ SEQUENCE 169 AA; 20167 MW; C1DEDFB1EF123898 CRC64;					

Query Match 14.0%; Score 7; DB 1; Length 169;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 LEATVDY 43  
|||||||  
Db 144 LEATVDY 150

RESULT 2  
Y9E3\_PSEAE  
ID Y9E3\_PSEAE STANDARD; PRT; 242 AA.  
AC P33663;

**THIS PAGE BLANK (USPTO)**



A:Residues: 1-390 <STO>  
 A:Cross-references: GB:AE004885; GB:AE004091; NID:g9950968; PIDN:AAG08108.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA4722  
 C:Superfamily: aspartate transaminase

Query Match 14.0%; Score 7; DB 2; Length 390;

Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 DAASVLE 38  
 |||||

Db 215 DAASVLE 221

RESULT 12

T48593 hypothetical protein T22N19.110 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T48593  
 R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,

submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z24490

A:Accession: T48593

A:Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-443 <BEV>  
 A:Cross-references: EMBL:AL163572

A:Experimental source: cultivar Columbia; BAC clone T22N19  
 C:Genetics:

A:Map position: 5

A:Introns: 23/3; 125/3; 196/3; 240/3  
 A:Note: T22N19.110

Query Match

Best Local Similarity 14.0%; Score 7; DB 2; Length 443;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERIK 13  
 |||||

Db 227 LRRERIK 233

RESULT 13

F64204 spermidine/putrescine transport ATP-binding protein potA homolog - Mycoplasma genitalium

C:Species: Mycoplasma genitalium  
 C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 17-Mar-2000

C:Accession: F64204  
 R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;

M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.  
 C.A.; Venter, J.C.

Science 270, 397-403, 1995  
 A:Title: The minimal gene complement of Mycoplasma genitalium.

A:Reference number: A64200; MUID:96026346  
 A:Accession: F64204

A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA

A:Residues: 1-486 <TIGR>  
 A:Cross-references: GB:U393683; GB:L43967; NID:g1045711; PID:g1045714; TIGR:MG042

A:Experimental source: strain G-37  
 C:Genetics:

A:Genetic code: SGC3

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
 C:Keywords: ATP

F:1-351/Domain: ATP-binding cassette homology #status atypical <ABCL>

Query Match

14.0%; Score 7; DB 2; Length 486;

Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KEKLRE 10  
 |||||

Db 100 KEKLRE 106

RESULT 14

A40283

3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17), calmodulin-dependent, 59K ca  
 C:Species: Bos primigenius taurus (cattle)

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 21-Jan-2000  
 C:Accession: A40283

R:Novack, J.P.; Charbonneau, H.; Bentley, J.K.; Walsh, K.A.; Beavo, J.A.  
 Biochemistry 30, 7940-7947, 1991

A:Title: Sequence comparison of the 63-, 61-, and 59-kDa calmodulin-dependent cyclic  
 A:Reference number: A40283; MUID:91329366

A:Accession: A40283

A:Molecule type: protein  
 A:Residues: 1-491 <NOV>

C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase 1B, calmodulin-dependent; 3',  
 C:Keywords: calmodulin binding; cardiac muscle; heart; phosphoric diester hydrolase

F:196-396/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNP>

Query Match

Best Local Similarity 14.0%; Score 7; DB 2; Length 491;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
 |||||

Db 43 AASVLEA 49

RESULT 15

T14783

hypothetical protein DKFZp586G0221.1 - human

C:Species: Homo sapiens (man)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jan-2000  
 C:Accession: T14783

R:Ottensmøller, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, August 1999

A:Reference number: Z18184

A:Accession: T14783

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-519 <OTT>

A:Cross-references: EMBL:AL110263

A:Experimental source: adult uterus; clone DKFZp586G0221

C:Genetics:

A:Note: DKFZp586G0221.1

C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase 1B, calmodulin-dependent; 3',  
 F:202-419/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNP>

Query Match

Best Local Similarity 14.0%; Score 7; DB 2; Length 519;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
 |||||

Db 43 AASVLEA 49

Search completed: September 15, 2001, 12:48:54  
 Job time: 224 sec

C:Accession: H85067  
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488  
A:Accession: H85067  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-250 <STO>  
A:Cross-references: GB:NC\_001268; NID:g7267300; PIDN:CAB81082.1; GSPDB:GN00140  
C:Genetics:  
A:Gene: AT4g05400  
A:Map position: 4  
C:Superfamily: Arabidopsis thaliana hypothetical protein F7J7.80

Query Match 14.0%; Score 7; DB 2; Length 250;  
Best Local Similarity 100.0%; Pred. No. 8.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKEKLR 8  
|||||||  
Db 244 SSKEKLR 250

RESULT 8  
B71312  
probable GTP-binding protein (era) - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 02-Feb-2001  
C:Accession: B71312  
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; MUID:98332770  
A:Accession: B71312  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-319 <COL>  
A:Cross-references: GB:AE001229; GB:AE000520; NID:g3322827; PIDN:AAC65525.1; PID:g332283  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0541  
C:Superfamily: ras transforming protein; translation elongation factor Tu homology  
C:Keywords: GTP binding; nucleotide binding; P-loop  
F.11-130/Domain: translation elongation factor Tu homology <ETU>  
F.17-24/Region: nucleotide-binding motif A (P-loop)  
F.127-130/Region: GTP-binding NKXD motif  
F.176-178/Region: GTP-binding SAK/L motif

Query Match 14.0%; Score 7; DB 2; Length 319;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLPLP 25  
|||||||  
Db 287 LRTLPLP 293

RESULT 9  
H72472  
hypothetical protein APE2425 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: H72472  
R:Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hailkawa, Y.; Jin-no, K.; Takah  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
awa Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:99310339

A:Accession: H72472  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-326 <KAW>  
A:Cross-references: DBJ:AP000064; NID:g5105945; PIDN:BAA81440.1; PID:g5106129  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE2425  
C:Superfamily: conserved hypothetical protein MJ1427

Query Match 14.0%; Score 7; DB 2; Length 326;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
|||||||  
Db 292 AASVLEA 298

RESULT 10  
E64188  
dipeptide transport ATP-binding protein dppF - Haemophilus influenzae (strain Rd KW20  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 20-Aug-1999  
C:Accession: E64188  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman  
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: E64188  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-327 <TIGR>  
A:Cross-references: GB:U32798; GB:L42023; NID:g1574110; PIDN:AAC22837.1; PID:g1574111  
A:Gene: dppF  
C:Function:  
A:Description: probably responsible for energy-coupling to the transport system  
C:Superfamily: inner membrane protein malk; ATP-binding cassette homology  
C:Keywords: ATP; binding protein-dependent transport system; membrane protein; nucleo  
F.37-237/Domain: ATP-binding cassette homology <ABC>  
F.54-62/Region: nucleotide-binding motif A (P-loop)  
F.181-185/Region: nucleotide-binding motif B

Query Match 14.0%; Score 7; DB 2; Length 327;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERIK 13  
|||||||  
Db 271 LRRERIK 277

RESULT 11  
D83057  
probable aminotransferase PA4722 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83057  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A:Reference number: A82950; MUID:20437337  
A:Accession: D83057  
A:Status: preliminary  
A:Molecule type: DNA



C:Function:

A:Description: confers resistance against Fusarium oxysporum

C:Superfamily: disease resistance protein RPS2; leucine-rich alpha-2-glycoprotein repeat

Query Match 16.0%; Score 8; DB 2; Length 1220;

Best Local Similarity 100.0%; Pred. No. 3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 EQLRTLTP 24

Db 554 EQLRTLTP 561

RESULT 3

T06404

resistance complex protein I2C-2 - tomato

C:Species: Lycopersicon esculentum (tomato)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 17-Nov-2000

C:Accession: T06404

R:Orl, N.; Eshed, Y.; Paran, I.; Presting, G.; Aviv, D.; Tanksley, S.; Zamir, D.; Fluhr,

Plant Cell 9, 521-532, 1997

A:Title: The I2C family from the wilt disease resistance locus I2 belongs to the nucleot

A:Reference number: Z15652; MUID:97290204

A:Accession: T06404

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1240 &lt;ORI&gt;

A:Cross-references: EMBL:AF0404879; NID:g2258316; PIDN:AAB63275.1; PID:g2258317

C:Genetics:

A:Gene: I2C-2

A:Map position: 11

C:Function:

A:Description: confers resistance against Fusarium oxysporum

C:Superfamily: disease resistance protein RPS2; leucine-rich alpha-2-glycoprotein repeat

Query Match

Best Local Similarity 16.0%; Score 8; DB 2; Length 1240;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 EQLRTLTP 24

Db 545 EQLRTLTP 552

RESULT 4

A64396

hypothetical protein MJ0769 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C:Accession: A64396

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; MUID:96337999

A:Accession: A64396

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-169 &lt;BUL&gt;

A:Cross-references: GB:U67522; GB:L77117; NID:g2826315; PIDN:AAB98774.1; PID:g1499589; T

C:Genetics:

A:Map position: REV690989-690480

Query Match

Best Local Similarity 14.0%; Score 7; DB 2; Length 169;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 LEATVDY 43

Db 144 LEATVDY 150

RESULT 5

A70527

hypothetical protein Rv0328 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C:Accession: A70527

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

Rajandream, M.A.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Nature 393, 537-544, 1998

A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: A70527

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-200 &lt;COL&gt;

A:Cross-references: GB:Z96800; GB:AL123456; NID:g3261800; PIDN:CAB09609.1; PID:g21939

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv0328

Query Match

Best Local Similarity 14.0%; Score 7; DB 2; Length 200;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 VLEATVD 42

Db 88 VLEATVD 94

RESULT 6

EB3077

conserved hypothetical protein PA4543 [imported] - Pseudomonas aeruginosa (strain PAO

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: EB3077

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa

A:Reference number: A82950; MUID:20437337

A:Accession: EB3077

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-242 &lt;STO&gt;

A:Cross-references: GB:AE004868; GB:AE004091; NID:g9950785; PIDN:ANG07931.1; GSPDB:GN

A:Experimental source: strain PAO1

C:Genetics:

A:Gene: PA4543

C:Superfamily: Streptomyces coelicolor hypothetical protein SC4A10.14c

Query Match

Best Local Similarity 14.0%; Score 7; DB 2; Length 242;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 VLEATVD 42

Db 130 VLEATVD 136

RESULT 7

H85067

hypothetical protein AT4g05400 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: September 15, 2001, 12:48:53 ; Search time 45.39 Seconds  
(without alignments)  
83.911 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_189  
Perfect score: 50  
Sequence: 1 HSSKEKLRRRIKYCCQLR.....NDAASVLEATVDYVKYIREK 50

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	16.0	703	G75638	nodulation protein
2	8	16.0	1220	T06403	resistance complex
3	8	16.0	1240	T06404	resistance complex
4	7	14.0	169	A64396	hypothetical prote
5	7	14.0	200	A70527	hypothetical prote
6	7	14.0	242	E83077	conserved hypothet
7	7	14.0	250	H85067	hypothetical prote
8	7	14.0	319	B71312	probable GTP-bindi
9	7	14.0	326	H72472	hypothetical prote
10	7	14.0	327	E64188	dipeptide transpor
11	7	14.0	390	D83057	probable aminotran
12	7	14.0	443	T48593	hypothetical prote
13	7	14.0	486	F64204	spermidine/putresc
14	7	14.0	491	A40283	3',5'-cyclic-nucle
15	7	14.0	519	T14783	hypothetical prote
16	7	14.0	530	A45334	3',5'-cyclic-nucle
17	7	14.0	586	C71695	ctp synthase (pyrG
18	7	14.0	748	T30634	hypothetical prote
19	7	14.0	770	A44337	kinesin-related pr
20	7	14.0	791	T20815	hypothetical prote
21	7	14.0	838	S49750	probable membrane
22	7	14.0	974	A71466	probable zinc meta
23	7	14.0	1270	T30339	dsRNA adenosine de
24	7	14.0	1341	S66835	probable membrane
25	6	12.0	59	ERBP1A	gene 1A protein
26	6	12.0	68	IABY3	proteinase A inhib
27	6	12.0	79	A82875	ribosomal protein
28	6	12.0	81	S61847	hrpx protein - pse
29	6	12.0	87	C82334	probable host fact

30	6	12.0	90	2	H69159	hypothetical prote
31	6	12.0	91	2	D64066	host factor I - Ha
32	6	12.0	93	2	T40184	conserved hypothet
33	6	12.0	94	2	T12719	hypothetical prote
34	6	12.0	109	2	F82067	probable anti-sigm
35	6	12.0	111	2	S37723	outer surface prot
36	6	12.0	122	2	A69812	hypothetical prote
37	6	12.0	134	2	J01403	acidic seminal flu
38	6	12.0	135	2	T19002	hypothetical prote
39	6	12.0	145	2	C71378	hypothetical prote
40	6	12.0	150	2	I38595	leukotriene-C4 syn
41	6	12.0	153	2	A29472	T-cell surface gly
42	6	12.0	153	2	S68406	vesicular integral
43	6	12.0	159	2	G84974	transcription elon
44	6	12.0	160	2	T24339	hypothetical prote
45	6	12.0	160	2	S52889	probable membrane

ALIGNMENTS

RESULT 1  
G75638  
nodulation protein-related protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
R:Accession: G75638  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Status: preliminary  
A:Accession: G75638  
A:Molecule type: DNA  
A:Residues: 1-703 <WHIT>  
A:Cross-references: GB:AE001827; NID:g6460959; PIDN:AAF12674.1; PID:g6460971; TIGR:DR  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DRC0037  
A:Map position: plasmid  
A:Genome: plasmid  
A:Note: plasmid Cpl

Query Match 16.0%; Score 8; DB 2; Length 703;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEAT 40  
|||||||  
DB 662 AASVLEAT 669

RESULT 2  
T06403  
resistance complex protein I2C-1 - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 17-Nov-2000  
C:Accession: T06403  
R:Ori, N.; Eshed, Y.; Paran, I.; Presting, G.; Aviv, D.; Tanksley, S.; Zamir, D.; Flu  
Plant Cell 9, 521-532, 1997  
A:Title: The I2C family from the wilt disease resistance locus I2 belongs to the nuclei  
A:Reference number: Z15652; MUID:97290204  
A:Accession: T06403  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1220 <ORI>  
A:Cross-references: EMBL:AF004878; NID:g2258314; PIDN:AAB63274.1; PID:g2258315  
C:Genetics:  
A:Gene: I2C-1  
A:Map position: 11



MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/297,510  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5602019and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-297-510-6

Query Match 14.0%; Score 7; DB 1; Length 530;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
Db 59 AASVLEA 65

## RESULT 15

US-08-479-532-6  
Sequence 6, Application US/08479532  
Patent No. 5776752  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
TITLE OF INVENTION: Phosphodiesterases  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,532  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/297,494  
FILING DATE:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:

NAME: No. 5776752and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-479-532-6

Query Match 14.0%; Score 7; DB 1; Length 530;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
Db 59 AASVLEA 65

Search completed: September 15, 2001, 12:47:54  
Job time: 184 sec

QY 33 AASVLEA 39  
| | | | |  
DB 43 AASVLEA 49

## RESULT 12

US-07-872-644-6  
; Sequence 6, Application US/07872644  
; Patent No. 5389527  
; GENERAL INFORMATION:  
; APPLICANT: Beavo, Joseph A.  
; APPLICANT: Bentley, Kelley  
; APPLICANT: Charbonneau, Harry  
; APPLICANT: Sonnenburg, William K.  
; TITLE OF INVENTION: DNA Encoding Mammalian  
; Phosphodiesterases  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 19920420  
; APPLICATION NUMBER: US/07/872,644  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/688,356  
; FILING DATE: 04-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5389527and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/30822  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 6  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-872-644-6

Query Match 14.0%; Score 7; DB 1; Length 530;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
| | | | |  
DB 59 AASVLEA 65

## RESULT 13

US-08-297-494-6  
; Sequence 6, Application US/08297494  
; Patent No. 5380771  
; GENERAL INFORMATION:  
; APPLICANT: Beavo, Joseph A.  
; APPLICANT: Bentley, Kelley  
; APPLICANT: Charbonneau, Harry

; APPLICANT: Sonnenburg, William K.  
; TITLE OF INVENTION: DNA Encoding Mammalian  
; Phosphodiesterases  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/297,494  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/688,356  
; FILING DATE: 04-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 558077land, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/30822  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 6  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-297-494-6

Query Match 14.0%; Score 7; DB 1; Length 530;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
| | | | |  
DB 59 AASVLEA 65

## RESULT 14

US-08-297-510-6  
; Sequence 6, Application US/08297510  
; Patent No. 5602019  
; GENERAL INFORMATION:  
; APPLICANT: Beavo, Joseph A.  
; APPLICANT: Bentley, Kelley  
; APPLICANT: Charbonneau, Harry  
; APPLICANT: Sonnenburg, William K.  
; TITLE OF INVENTION: DNA Encoding Mammalian  
; Phosphodiesterases  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
COMPUTER READABLE FORM:

;; PRIOR APPLICATION DATA:  
;; FILING DATE: 08/297,494  
;; APPLICATION NUMBER: US 07/688,356  
;; FILING DATE: 04-APR-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. 5800987and, Greta E.  
;; REGISTRATION NUMBER: 35,302  
;; REFERENCE/DOCKET NUMBER: 27866/30822  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (312) 346-5750  
;; TELEFAX: (312) 984-9740  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 17:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 514 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-455-525-17

Query Match 14.0%; Score 7; DB 1; Length 514;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
|||||||  
Db 43 AASVLEA 49

## RESULT 10

US-09-139-491-17  
;; Sequence 17, Application US/09139491  
;; Patent No. 6015677  
;; GENERAL INFORMATION:  
;; APPLICANT: Beavo, Joseph A.  
;; APPLICANT: Bentley, Kelley  
;; APPLICANT: Charbonneau, Harry  
;; APPLICANT: Sonnenburg, William K.  
;; TITLE OF INVENTION: DNA Encoding Mammalian  
;; NUMBER OF SEQUENCES: 58  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
;; STREET: Two First National Plaza, 20 South Clark  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60603  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION NUMBER: US/09/139,491  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/455,525  
;; FILING DATE: 31-MAY-1995  
;; APPLICATION NUMBER: 08/297,494  
;; FILING DATE:  
;; APPLICATION NUMBER: US 07/688,356  
;; FILING DATE: 04-APR-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. 6015677and, Greta E.  
;; REGISTRATION NUMBER: 35,302  
;; REFERENCE/DOCKET NUMBER: 27866/30822  
;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (312) 346-5750  
;; TELEFAX: (312) 984-9740  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 17:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 514 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-09-139-491-17

Query Match 14.0%; Score 7; DB 3; Length 514;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
|||||||  
Db 43 AASVLEA 49

## RESULT 11

PCT-US92-03222-17  
;; Sequence 17, Application PC/TUS9203222  
;; GENERAL INFORMATION:  
;; APPLICANT: Beavo, Joseph A.  
;; APPLICANT: Bentley, Kelley  
;; APPLICANT: Charbonneau, Harry  
;; APPLICANT: Sonnenburg, William K.  
;; TITLE OF INVENTION: DNA Encoding Mammalian  
;; NUMBER OF SEQUENCES: 58  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
;; STREET: Two First National Plaza, 20 South Clark  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60603  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION NUMBER: PCT/US92/03222  
;; FILING DATE: 19920420  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/688,356  
;; FILING DATE: 04-APR-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Noland, Greta E.  
;; REGISTRATION NUMBER: 35,302  
;; REFERENCE/DOCKET NUMBER: 27866/30822  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (312) 346-5750  
;; TELEFAX: (312) 984-9740  
;; TELEX: 25-3856  
;; INFORMATION FOR SEQ ID NO: 17:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 514 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
PCT-US92-03222-17

Query Match 14.0%; Score 7; DB 5; Length 514;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/297,494  
FILING DATE:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5776752and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-479-532-17

Query Match 14.0%; Score 7; DB 1; Length 514;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 AASVLEA 39  
Db 43 AASVLEA 49

RESULT 8  
US-08-455-526-17  
Sequence 17, Application US/08455526  
Patent No. 5789553  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago

STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/297,494  
FILING DATE: 29-AUG-1994  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5789553and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-455-526-17

Query Match 14.0%; Score 7; DB 1; Length 514;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 AASVLEA 39  
Db 43 AASVLEA 49

RESULT 9  
US-08-455-525-17  
Sequence 17, Application US/08455525  
Patent No. 5800987  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Bicknell  
STREET: Two First National Plaza, 20 South Clark  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,525  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 435

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-872-644-17

Query Match 14.0%; Score 7; DB 1; Length 514;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
Db 43 AASVLEA 49

RESULT 5  
US-08-297-494-17  
; Sequence 17, Application US/08297494  
; Patent No. 5580771  
; GENERAL INFORMATION:  
; APPLICANT: Beavo, Joseph A.  
; APPLICANT: Bentley, Kelley  
; APPLICANT: Charbonneau, Harry  
; APPLICANT: Sonnenburg, William K.  
; TITLE OF INVENTION: DNA Encoding Mammalian  
; TITLE OF INVENTION: Phosphodiesterases  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; STREET: Street  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/297,494  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/688,356  
; FILING DATE: 04-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5580771and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/30822  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 514 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-297-494-17

Query Match 14.0%; Score 7; DB 1; Length 514;

Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
Db 43 AASVLEA 49

RESULT 6  
US-08-297-510-17  
; Sequence 17, Application US/08297510  
; Patent No. 5602019  
; GENERAL INFORMATION:  
; APPLICANT: Beavo, Joseph A.  
; APPLICANT: Bentley, Kelley  
; APPLICANT: Charbonneau, Harry  
; APPLICANT: Sonnenburg, William K.  
; TITLE OF INVENTION: DNA Encoding Mammalian  
; TITLE OF INVENTION: Phosphodiesterases  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
; ADDRESSEE: Bicknell  
; STREET: Two First National Plaza, 20 South Clark  
; STREET: Street  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/297,510  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/688,356  
; FILING DATE: 04-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5602019and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/30822  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 346-5750  
; TELEFAX: (312) 984-9740  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 514 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-297-510-17

Query Match 14.0%; Score 7; DB 1; Length 514;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVLEA 39  
Db 43 AASVLEA 49

RESULT 7  
US-08-479-532-17  
; Sequence 17, Application US/08479532  
; Patent No. 5776752  
; GENERAL INFORMATION:  
; APPLICANT: Beavo, Joseph A.



Db 554 EQLRTLLP 561

RESULT 2

US-08-930-996A-4

; Sequence 4, Application US/08930996A

; Patent No. 6100449

; GENERAL INFORMATION:

; APPLICANT: FLUHR, Robert

; APPLICANT: ESHED, Yuval

; APPLICANT: ORI, Naomi

; APPLICANT: PARAN, Ilan

; APPLICANT: ZAMIR, Daniel

; TITLE OF INVENTION: A GENE FAMILY FROM THE I2 FUSARIUM RESISTANCE

; TITLE OF INVENTION: LOCUS OF TOMATO AND USE THEREOF FOR TRANSFORMATION AND

; TITLE OF INVENTION: SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK

; STREET: 419 Seventh Street, N.W., Suite 300

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/930.996A

; FILING DATE: 09-DEC-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/05272

; FILING DATE: 15-APR-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: IL 113,373

; FILING DATE: 13-APR-1995

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1240 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-930-996A-4

Query Match 16.0%; Score 8; DB 3; Length 1240;

Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 EQLRTLLP 24

Db 545 EQLRTLLP 552

RESULT 3

US-08-630-916A-76

; Sequence 76, Application US/08630916A

; Patent No. 601137

; GENERAL INFORMATION:

; APPLICANT: Pirozzi, Gregorio

; APPLICANT: Kay, Brian K.

; APPLICANT: Fowkes, Dana M.

; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL

; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME

; NUMBER OF SEQUENCES: 124

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: United States

US-08-630-916A-76

Query Match 14.0%; Score 7; DB 3; Length 52;

Best Local Similarity 100.0%; Pred. No. 0.98;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSKEKLR 8

Db 30 SSKEKLR 36

RESULT 4

US-07-872-644-17

; Sequence 17, Application US/07872644

; Patent No. 5389527

; GENERAL INFORMATION:

; APPLICANT: Beavo, Joseph A.

; APPLICANT: Bentley, Kelley

; APPLICANT: Charbonneau, Harry

; APPLICANT: Sonnenburg, William K.

; TITLE OF INVENTION: DNA Encoding Mammalian

; TITLE OF INVENTION: Phosphodiesterases

; NUMBER OF SEQUENCES: 58

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

; ADDRESSEE: Bicknell

; STREET: Two First National Plaza, 20 South Clark

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60603

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/872.644

; FILING DATE: 19920420

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/688,356

; FILING DATE: 04-APR-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 5389527and, Greta E.

; REGISTRATION NUMBER: 35,302

; REFERENCE/DOCKET NUMBER: 27866/30822

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: September 15, 2001, 12:47:53 ; Search time 35.36 Seconds  
(without alignments)  
29.115 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_189  
Perfect score: 50  
Sequence: 1 HSSKELRRRIKYCCQLR.....NDAASVLEATVDYVKYIREK 50

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 197339 seqs, 20590346 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents\_AA:  
1: /cgnl\_7/ptodata/1/iaa/5A.COMB.pap:\*  
2: /cgnl\_7/ptodata/1/iaa/5B.COMB.pap:\*  
3: /cgnl\_7/ptodata/1/iaa/6A.COMB.pap:\*  
4: /cgnl\_7/ptodata/1/iaa/6B.COMB.pap:\*  
5: /cgnl\_7/ptodata/1/iaa/PCTUS.COMB.pap:\*  
6: /cgnl\_7/ptodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES:

Result No.	Score	Query Match	Length	ID	Description
1	8	16.0	1220	3	US-08-930-996A-2
2	8	16.0	1240	3	US-08-930-996A-4
3	7	14.0	52	3	US-08-630-916A-76
4	7	14.0	514	1	US-07-872-644-17
5	7	14.0	514	1	US-08-297-494-17
6	7	14.0	514	1	US-08-297-510-17
7	7	14.0	514	1	US-08-479-532-17
8	7	14.0	514	1	US-08-455-526-17
9	7	14.0	514	1	US-08-455-525-17
10	7	14.0	514	3	US-09-139-491-17
11	7	14.0	514	5	PCT-US92-03222-17
12	7	14.0	530	1	US-07-872-644-6
13	7	14.0	530	1	US-08-297-494-6
14	7	14.0	530	1	US-08-297-510-6
15	7	14.0	530	1	US-08-479-532-6
16	7	14.0	530	1	US-08-455-526-6
17	7	14.0	530	1	US-08-455-525-6
18	7	14.0	530	3	US-09-139-491-6
19	7	14.0	530	5	PCT-US92-03222-6
20	7	14.0	535	1	US-07-872-644-49
21	7	14.0	535	1	US-08-297-494-49
22	7	14.0	535	1	US-08-297-510-49
23	7	14.0	535	1	US-08-479-532-49
24	7	14.0	535	1	US-08-455-526-49
25	7	14.0	535	1	US-08-455-525-49
26	7	14.0	535	3	US-09-139-491-49
27	7	14.0	535	5	PCT-US92-03222-49

28	6	12.0	15	4	US-08-986-837-8	Sequence 8, Appl1
29	6	12.0	27	1	US-08-447-702-1	Sequence 1, Appl1
30	6	12.0	27	1	US-08-465-615-1	Sequence 1, Appl1
31	6	12.0	79	2	US-08-611-510-6	Sequence 6, Appl1
32	6	12.0	147	4	US-08-986-837-2	Sequence 2, Appl1
33	6	12.0	150	2	US-08-254-354-2	Sequence 2, Appl1
34	6	12.0	150	5	PCT-US95-06137-2	Sequence 2, Appl1
35	6	12.0	153	2	US-08-695-736-4	Sequence 4, Appl1
36	6	12.0	268	3	US-08-871-483-11	Sequence 11, Appl1
37	6	12.0	273	4	US-08-235-836C-98	Sequence 98, Appl1
38	6	12.0	273	4	US-08-235-836C-101	Sequence 101, Appl1
39	6	12.0	274	2	US-08-441-857-2	Sequence 2, Appl1
40	6	12.0	274	3	US-08-193-159-2	Sequence 2, Appl1
41	6	12.0	274	4	US-08-235-836C-9	Sequence 9, Appl1
42	6	12.0	274	4	US-08-235-836C-86	Sequence 86, Appl1
43	6	12.0	274	4	US-08-235-836C-138	Sequence 138, Appl1
44	6	12.0	289	2	US-08-741-437-1	Sequence 1, Appl1
45	6	12.0	289	2	US-08-741-437-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1  
US-08-930-996A-2  
; Sequence 2, Application US/08930996A  
; Patent No. 6100449  
; GENERAL INFORMATION:  
; APPLICANT: FLUHR, Robert  
; APPLICANT: ESHED, Yuval  
; APPLICANT: ORI, Naomi  
; APPLICANT: PARAN, Ilan  
; APPLICANT: ZAMIR, Daniel  
; TITLE OF INVENTION: A GENE FAMILY FROM THE I2 FUSARIUM RESISTANCE  
; TITLE OF INVENTION: LOCUS OF TOMATO AND USE THEREOF FOR TRANSFORMATION AND  
; TITLE OF INVENTION: SELECTIVE BREEDING OF TOMATO AND RELATED PLANTS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/930,996A  
; FILING DATE: 09-DEC-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/05272  
; FILING DATE: 15-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: IL 113,373  
; FILING DATE: 13-APR-1995  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1220 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-930-996A-2

Query Match 16.0%; Score 8; DB 3; Length 1220;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 17 EOLRTLLP 24  
|||||||



PA (UNIW ) UNIV WASHINGTON.  
 XX Beavo JA, Bentley JK, Charbonneau H, Sonnenburg WK;  
 PI WPI; 1992-382051/46.  
 XX N-PSDB; AAQ30172.  
 DR  
 XX  
 PT New DNA encoding mammalian cyclic nucleotide phospho-di-esterase  
 PT - and derived vectors and host cells, useful for screening cpds.  
 PT for inhibitory or activating activity  
 XX  
 PS Example 2; Page 24; 133pp; English.  
 XX  
 CC mRNA from bovine heart was used conventionally to generate first  
 CC strand cDNA and the prod. subjected to PCR amplification in the  
 CC presence of redundant sense and antisense oligonucleotide primers  
 CC corresp. to a partial sequence of a bovine heart 59 kD calcium/  
 CC calmodulin-stimulated cyclic nucleotide phosphodiesterase  
 CC polypeptide (Cam PDE) and the bovine brain 61 kD Cam PDE to create  
 CC a predicted 54 bp prod. with minimum overlap between the 59 and 61 kD  
 CC isozymes. The PCR prod. was used to probe a bovine lung cDNA  
 CC library to isolate a clone, p59KCAMPDE-2, contg. the complete coding  
 CC sequence of the putative 59 kD Cam PDE. The protein prod. has an  
 CC estimated mol. wt. of ca. 59 kD and is nearly identical to the 61 kD  
 CC Cam PDE amino acid sequence except for the amino terminal 18 residues.  
 XX  
 SQ Sequence 514 AA;

Query Match 14.0%; Score 7; DB 13; Length 514;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 33 AASVLEA 39  
 Db 43 aasvlea 49  
 |||||

RESULT 15  
 AAR69715  
 ID AAR69715 standard; Protein; 514 AA.  
 XX  
 AC AAR69715;  
 XX  
 DT 11-OCT-1995 (first entry)  
 XX  
 DE Cyclic-GMP stimulated nucleotide PDE clone p59KCAMPDE-2.  
 XX  
 KW Cyclic-GMP stimulated nucleotide phospho-diesterase;  
 KW bovine lung; hormones; neurotransmitters; transmission regulation;  
 KW antibodies; enzyme purification; clone p59KCAMPDE-2.  
 XX  
 OS Bos taurus.  
 XX  
 PN US5389527-A.  
 XX  
 PD 14-FEB-1995.  
 XX  
 PF 19-APR-1991; 91US-0688356.  
 XX  
 PR 19-APR-1991; 91US-0688356.  
 PR 20-APR-1992; 92US-0872644.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Beavo JA, Charbonneau H, Sonnenburg WK;  
 XX  
 DR WPI; 1995-090205/12.  
 DR N-PSDB; AAQ83964.  
 XX  
 PT New nucleic acid encoding cyclic-GMP stimulated nucleotide  
 PT phospho-di-esterase - and related vectors and transformed cells,  
 PT useful for screening cpds. for phospho-di-esterase modulating

PT activity  
 XX  
 PS Example 2; Columns 43-48; 69pp; English.  
 XX  
 CC AAQ83964 encodes AAR69715 bovine lung cyclic-GMP stimulated nucleotide  
 CC phospho-diesterase (Cam PDE) clone p59KCAMPDE-2. Eukaryotic cells  
 CC that express Cam PDE can be used to screen cpds. for the ability to  
 CC modulate Cam PDE activity. Cam PDEs are involved in regulating  
 CC the transmission of information from hormones, neurotransmitters  
 CC or other systems that use cyclic nucleotides as messengers.  
 CC Antibodies raised against Cam PDE can be used for enzyme purifcn..  
 CC or determination.  
 XX

SQ Sequence 514 AA;  
 Query Match 14.0%; Score 7; DB 16; Length 514;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 33 AASVLEA 39  
 Db 43 aasvlea 49  
 |||||

Search completed: September 15, 2001, 12:47:12  
 Job time: 162 sec

```

PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 14.0%; Score 7; DB 21; Length 443;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERIK 13
DB 227 Lrrerik 233

RESULT 13
AAY37305
ID AAY37305 standard; Protein; 477 AA.
XX AC
XX AAY37305;
XX 07-OCT-1999 (first entry)
DT
DE Amino acid sequence of a Chlamydia trachomatis protein.
XX KW
XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;

```

```

KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
XX bartholinitis; pneumopathy; venereal lymphogranulomatosis.
OS Chlamydia trachomatis.
XX WO9928475-A2.
XX 10-JUN-1999.
XX 27-NOV-1998; 98WO-IB01939.
XX 04-NOV-1998; 98US-0107077.
PR 18-NOV-1997; 97FR-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX (GEST ) GENSET.
XX Griffais R;
XX WPI; 1999-371125/31.
XX Genome sequence of Chlamydia trachomatis
PS Disclosure; Page 1042; 1755pp; English.
XX
CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
CC perihepatitis, bartholinitis; pneumopathy in breast feeding infants;
CC and venereal lymphogranulomatosis. The polypeptides of the invention
CC may be of use in treating these diseases.
XX
SQ Sequence 477 AA:

Query Match 14.0%; Score 7; DB 20; Length 477;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 YVKYIRE 49
DB 203 yvkYire 209

RESULT 14
AAR28401
ID AAR28401 standard; Protein; 514 AA.
XX AC
XX AAR28401;
XX 19-MAR-1993 (first entry)
DT
DE Bovine lung 59 kD Cam PDE.
XX
KW Calcium/calmodulin; stimulated; cyclic; nucleotide;
KW phosphodiesterase.
XX Bos taurus.
XX OS
XX WO9218541-A.
XX 29-OCT-1992.
XX 20-APR-1992; 92WO-US03222.
XX 19-APR-1991; 91US-0688356.
XX

```

AC AAG15756;  
XX 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 16133.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN  
XX  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136382.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139482.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.

PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 23-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 13-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 14.0%; Score 7; DB 21; Length 341;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRERIK 13  
|||||  
Db 125 lrrerik 131

RESULT 12  
AAG15756  
ID AAG15756 standard; Protein: 443 AA.  
XX

```
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 14.0%; Score 7; DB 21; Length 291;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 LRRERIK 13
|||||||
```

Db 75 lrrerik 81

## RESULT 11

AAG15757

ID AAG15757 standard; Protein; 341 AA.

XX

AC AAG15757;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 16134.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR

PR 05-MAR-1999; 99US-0123180.

PR

PR 09-MAR-1999; 99US-0123548.

PR

PR 23-MAR-1999; 99US-0125788.

PR

PR 25-MAR-1999; 99US-0126264.

PR

PR 29-MAR-1999; 99US-0126785.

PR

PR 01-APR-1999; 99US-0127462.

PR

PR 06-APR-1999; 99US-0128234.

PR

PR 08-APR-1999; 99US-0128714.

PR

PR 16-APR-1999; 99US-0129845.

PR

PR 19-APR-1999; 99US-0130077.

PR

PR 21-APR-1999; 99US-0130449.

PR

PR 23-APR-1999; 99US-0130510.

PR

PR 23-APR-1999; 99US-0130891.

PR

PR 28-APR-1999; 99US-0131449.

PR

PR 30-APR-1999; 99US-0132048.

PR

PR 30-APR-1999; 99US-0132407.

PR

PR 04-MAY-1999; 99US-0132484.

PR

PR 05-MAY-1999; 99US-0132485.

PR

PR 06-MAY-1999; 99US-0132486.

PR

PR 06-MAY-1999; 99US-0132487.

PR

PR 07-MAY-1999; 99US-0132863.

PR

PR 11-MAY-1999; 99US-0134256.

PR

PR 14-MAY-1999; 99US-0134218.

PR

PR 14-MAY-1999; 99US-0134219.

PR

PR 14-MAY-1999; 99US-0134221.

PR

PR 14-MAY-1999; 99US-0134370.

PR

PR 18-MAY-1999; 99US-0134768.

PR

PR 19-MAY-1999; 99US-0134941.

PR

PR 20-MAY-1999; 99US-0135124.

PR

PR 21-MAY-1999; 99US-0135353.

PR

PR 24-MAY-1999; 99US-0135629.

PR

PR 25-MAY-1999; 99US-0136021.

PR

PR 27-MAY-1999; 99US-0136392.

PR

PR 28-MAY-1999; 99US-0136782.

PR

PR 01-JUN-1999; 99US-0137222.

PR

PR 03-JUN-1999; 99US-0137528.

PR

PR 04-JUN-1999; 99US-0137502.

PR

PR 07-JUN-1999; 99US-0137724.

PR

PR 08-JUN-1999; 99US-0138094.

PR

PR 10-JUN-1999; 99US-0138540.

PR

PR 10-JUN-1999; 99US-0138847.

PR

PR 14-JUN-1999; 99US-0139119.

PR

PR 16-JUN-1999; 99US-0139452.

PR

PR 16-JUN-1999; 99US-0139453.

PR

PR 17-JUN-1999; 99US-0139492.

PR

PR 18-JUN-1999; 99US-0139454.

PR

PR 18-JUN-1999; 99US-0139455.

PR



```
Query Match      14.0%; Score 7; DB 21; Length 250;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 SKEKLR 8
Db  244 sakeklr 250
      |||||
      |||||

RESULT 10
AAG15758
ID  AAG15758 standard; Protein; 291 AA.
XX  AC
XX  AAG15758;
DT  17-OCT-2000 (first entry)
XX  XX
XX  Arabidopsis thaliana protein fragment SEQ ID NO: 16135.
DE  XX
KW  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
KW  termination sequence.
XX  XX
OS  Arabidopsis thaliana.
XX  XX
PN  EP1033405-A2.
XX  XX
PD  06-SEP-2000.
XX  XX
PF  25-FEB-2000; 2000EP-0301439.
XX  XX
PR  25-FEB-1999; 99US-0121825.
PR  05-MAR-1999; 99US-0123180.
PR  09-MAR-1999; 99US-0123548.
PR  23-MAR-1999; 99US-0125788.
PR  25-MAR-1999; 99US-0126264.
PR  29-MAR-1999; 99US-0126785.
PR  01-APR-1999; 99US-0127462.
PR  06-APR-1999; 99US-0128234.
PR  08-APR-1999; 99US-0128714.
PR  16-APR-1999; 99US-0129845.
PR  19-APR-1999; 99US-0130077.
PR  21-APR-1999; 99US-0130449.
PR  23-APR-1999; 99US-0130510.
PR  23-APR-1999; 99US-0130891.
PR  28-APR-1999; 99US-0131449.
PR  30-APR-1999; 99US-0132048.
PR  30-APR-1999; 99US-0132407.
PR  04-MAY-1999; 99US-0132484.
PR  05-MAY-1999; 99US-0132485.
PR  06-MAY-1999; 99US-0132486.
PR  06-MAY-1999; 99US-0132487.
PR  07-MAY-1999; 99US-0132863.
PR  11-MAY-1999; 99US-0134256.
PR  14-MAY-1999; 99US-0134218.
PR  14-MAY-1999; 99US-0134219.
PR  14-MAY-1999; 99US-0134221.
PR  14-MAY-1999; 99US-0134370.
PR  18-MAY-1999; 99US-0134768.
PR  19-MAY-1999; 99US-0134941.
PR  20-MAY-1999; 99US-0135124.
PR  21-MAY-1999; 99US-0135353.
PR  24-MAY-1999; 99US-0135629.
PR  25-MAY-1999; 99US-0136021.
PR  27-MAY-1999; 99US-0136392.
PR  28-MAY-1999; 99US-0136782.
PR  01-JUN-1999; 99US-0137222.
PR  03-JUN-1999; 99US-0137528.
PR  04-JUN-1999; 99US-0137502.
PR  07-JUN-1999; 99US-0137724.
PR  08-JUN-1999; 99US-0138094.
PR  10-JUN-1999; 99US-0138540.
PR  10-JUN-1999; 99US-0138847.
PR  14-JUN-1999; 99US-0139119.
PR  16-JUN-1999; 99US-0139452.
PR  16-JUN-1999; 99US-0139453.
PR  17-JUN-1999; 99US-0139492.
PR  18-JUN-1999; 99US-0139454.
PR  18-JUN-1999; 99US-0139455.
PR  18-JUN-1999; 99US-0139456.
PR  18-JUN-1999; 99US-0139457.
PR  18-JUN-1999; 99US-0139458.
PR  18-JUN-1999; 99US-0139459.
PR  18-JUN-1999; 99US-0139460.
PR  18-JUN-1999; 99US-0139461.
PR  18-JUN-1999; 99US-0139462.
PR  18-JUN-1999; 99US-0139463.
PR  18-JUN-1999; 99US-0139750.
PR  18-JUN-1999; 99US-0139763.
PR  21-JUN-1999; 99US-0139817.
PR  22-JUN-1999; 99US-0139899.
PR  23-JUN-1999; 99US-0140353.
PR  23-JUN-1999; 99US-0140354.
PR  24-JUN-1999; 99US-0140695.
PR  28-JUN-1999; 99US-0140823.
PR  29-JUN-1999; 99US-0140991.
PR  30-JUN-1999; 99US-0141287.
PR  01-JUL-1999; 99US-0141842.
PR  01-JUL-1999; 99US-0142154.
PR  02-JUL-1999; 99US-0142055.
PR  06-JUL-1999; 99US-0142390.
PR  08-JUL-1999; 99US-0142803.
PR  09-JUL-1999; 99US-0142920.
PR  12-JUL-1999; 99US-0142977.
PR  13-JUL-1999; 99US-0143542.
PR  14-JUL-1999; 99US-0143624.
PR  15-JUL-1999; 99US-0144005.
PR  16-JUL-1999; 99US-0144085.
PR  16-JUL-1999; 99US-0144086.
PR  19-JUL-1999; 99US-0144325.
PR  19-JUL-1999; 99US-0144331.
PR  19-JUL-1999; 99US-0144332.
PR  19-JUL-1999; 99US-0144333.
PR  19-JUL-1999; 99US-0144334.
PR  19-JUL-1999; 99US-0144335.
PR  20-JUL-1999; 99US-0144352.
PR  20-JUL-1999; 99US-0144632.
PR  20-JUL-1999; 99US-0144884.
PR  21-JUL-1999; 99US-0144814.
PR  21-JUL-1999; 99US-0145086.
PR  21-JUL-1999; 99US-0145088.
PR  22-JUL-1999; 99US-0145085.
PR  22-JUL-1999; 99US-0145087.
PR  22-JUL-1999; 99US-0145089.
PR  22-JUL-1999; 99US-0145192.
PR  23-JUL-1999; 99US-0145145.
PR  23-JUL-1999; 99US-0145218.
PR  23-JUL-1999; 99US-0145224.
PR  26-JUL-1999; 99US-0145276.
PR  27-JUL-1999; 99US-0145913.
PR  27-JUL-1999; 99US-0145918.
PR  27-JUL-1999; 99US-0145919.
PR  28-JUL-1999; 99US-0145951.
PR  02-AUG-1999; 99US-0146386.
PR  02-AUG-1999; 99US-0146388.
PR  02-AUG-1999; 99US-0146389.
PR  03-AUG-1999; 99US-0147038.
PR  04-AUG-1999; 99US-0147204.
PR  05-AUG-1999; 99US-0147302.
PR  05-AUG-1999; 99US-0147192.
PR  05-AUG-1999; 99US-0147260.
PR  06-AUG-1999; 99US-0147303.
PR  06-AUG-1999; 99US-0147416.
PR  09-AUG-1999; 99US-0147493.
PR  09-AUG-1999; 99US-0147935.
```

PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0143977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160880.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 14.0%; Score 7; DB 21; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 7.9;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKELR 8  
 |||||  
 Db 129 sskelr 135

RESULT 8  
 AAY29197  
 ID AAY29197 standard; Protein; 170 AA.  
 XX  
 AC  
 XX  
 XX  
 DT 25-OCT-1999 (first entry)  
 XX  
 XX

DE Amino acid sequence of a virulence factor encoded by ORF26844c.  
 XX  
 KW Human pathogen: virulence polypeptide; virulence factor;  
 KW pathogenic infection; Pseudomonas aeruginosa infection.  
 XX  
 OS Pseudomonas aeruginosa.  
 XX  
 PN W09927129-A1.  
 XX  
 PD 03-JUN-1999.  
 XX  
 XX

PF 25-NOV-1998; 98WO-US25247.  
 XX  
 PR 25-NOV-1997; 97US-0066517.  
 XX  
 XX (GEO) GEN HOSPITAL CORP.  
 XX  
 XX Ausubel F, Cao H, Drenkard E, Goodman HM, Mahajan-Miklos S;  
 PI Rahme LG, Tan M, Tsongalis J;  
 XX  
 DR WPI; 1999-357851/30.  
 XX  
 PT Virulence factors useful in developing disease treatments  
 XX  
 XX Disclosure; Fig 4; 228pp; English.

XX  
 CC The present sequence represents a Pseudomonas aeruginosa polypeptide  
 CC sequence. P. aeruginosa is an opportunistic human pathogen present in  
 CC soil water and plants. The specification describes virulence polypeptides  
 CC and nucleic acid sequence encoding such polypeptides. These sequences  
 CC can be used to identify a compound which is capable of decreasing the  
 CC expression of a pathogenic virulence factor. Compounds that inhibit  
 CC the expression or activity of virulence factor polypeptides can be  
 CC used to treat pathogenic infections, especially where the infection  
 CC is a P. aeruginosa infection.

CC note: the sequences given in the specification were poorly legible, and  
 CC in some instances assumptions were made as to the identity of the  
 CC residue; it is therefore possible that the sequence given below is  
 CC not entirely correct.

SQ Sequence 170 AA;

Query Match 14.0%; Score 7; DB 20; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 9.7;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VLEATVD 42  
 |||||  
 Db 58 vleatvd 64

RESULT 9  
 AAG43790  
 ID AAG43790 standard; Protein; 250 AA.  
 XX  
 AC  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 54774.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 XX

PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.

PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134768.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0135782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147493.  
PR 10-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 12-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.

```

XX 05-DEC-1997 (first entry)
DT
XX
XX Tomato immunity 2 (fungal resistance) gene product.
DE
XX
XX Tomato; transgenic plant; disease resistance; wilt inducing fungi;
KW amplified fragment length polymorphism; AFLP; Immunity 2; I-2 gene;
KW Fusarium oxysporum; genetic engineering.
XX
XX Lycopersicon esculentum.
OS
XX
XX WO9706259-A2.
PN
XX
XX 20-FEB-1997.
PD
XX
XX 06-AUG-1996; 96WO-EP03480.
PF
XX
XX 07-AUG-1995; 95EP-0401849.
PR
XX
XX (KEYG-) KEYGENE NV.
PA
XX
XX Simons G, Vos P, Zabeau M;
PI
XX
XX WPI; 1997-154265/14.
DR
XX
XX N-PSDB; AAT79882.
DR
XX
XX New immunity-2 resistance gene, imparting pathogen resistance to
PT plants - used to transform plants, esp. to protect against
PT wilt-inducing fungi
PT
XX
XX Claim 41; Fig 6a-d; 6lpp; English.
PS
XX
XX AA25157 shows the product of the tomato immunity 2 (I-2) gene. This
CC gene was used to produce transgenic plants that are resistant to
CC wilt-inducing fungi, e.g. Fusarium 2, especially F. oxysporum f.sp.
CC lycopersici race 2. Plants that may be transformed include tomato,
CC melon, tobacco, Arabidopsis, aubergine and potato.
XX
XX Sequence 1266 AA;
SQ
Query Match 16.0%; Score 8; DB 18; Length 1266;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 17 EQLRTLTP 24
Db 547 eqlrtltp 554
|||||
RESULT 6
AAB07754
ID AAB07754 standard; Protein; 1266 AA.
XX
XX AAB07754;
AC
XX
XX 07-NOV-2000 (first entry)
DT
XX
XX Amino acid sequence of the protein encoded by the I-2 resistance gene.
DE
XX
XX Regulatory activity; transcription; I-2 resistance gene; tomato;
KW egg plant; potato; melon; tobacco; Arabidopsis; plant pathogen; fungi;
KW tissue-specific.
XX
XX Fusarium oxysporum.
OS
XX
XX EP1024196-A1.
PN
XX
XX 02-AUG-2000.
PD
XX
XX 29-JAN-1999; 99EP-0400212.
XX
XX 29-JAN-1999; 99EP-0400212.
PR

```

```

XX (KEYG-) KEYGENE NV.
PA
XX
XX Haring MA, Cornelissen BJC, Mes JJ, Simons AFM;
PI
XX
XX WPI; 2000-516034/47.
DR
XX
XX N-PSDB; AAA59332.
DR
XX
XX New I-2 resistance gene tissue-specific regulatory sequence useful in
PT plant resistance mechanisms against plant pathogens such as fungi -
PT
XX
XX Disclosure; Page 26-31; 47pp; English.
PS
XX
XX The present sequence represents I-2 resistance protein. The specification
CC describes nucleotide sequences which have a regulatory activity on the
CC transcription of the I-2 resistance gene in plant host cells.
CC The transgenic plants, especially tomato, egg plant, potato, melon,
CC tobacco and Arabidopsis, are capable of expressing a gene mediating
CC resistance to a plant pathogen, such as fungi, in a tissue-specific
CC manner. The plant is capable of preventing infection by a plant
CC pathogen, such as fungi. Inserting the regulatory activity polynucleotide
CC into plant cell genomes is useful for providing plants with reduced
CC susceptibility to plant pathogens, especially for protecting plants
CC in cultivation.
XX
XX Sequence 1266 AA;
SQ
Query Match 16.0%; Score 8; DB 21; Length 1266;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 17 EQLRTLTP 24
Db 547 eqlrtltp 554
|||||
RESULT 7
AAG43791
ID AAG43791 standard; Protein; 135 AA.
XX
XX AAG43791;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 54775.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR
XX
XX 05-MAR-1999; 99US-0123180.
PR
XX
XX 09-MAR-1999; 99US-0123548.
PR
XX
XX 23-MAR-1999; 99US-0125788.
PR
XX
XX 25-MAR-1999; 99US-0126264.
PR
XX
XX 29-MAR-1999; 99US-0126785.
PR
XX
XX 01-APR-1999; 99US-0127462.
PR
XX
XX 06-APR-1999; 99US-0128234.
PR
XX
XX 08-APR-1999; 99US-0128714.
PR
XX
XX 16-APR-1999; 99US-0129845.
PR
XX
XX 19-APR-1999; 99US-0130077.
PR
XX
XX 21-APR-1999; 99US-0130449.
PR
XX
XX 23-APR-1999; 99US-0130510.
PR
XX
XX 23-APR-1999; 99US-0130891.
PR

```

```

FH Domain 646..1220
FT /note= "Leucine-rich repeat region"
FT
FT
PN WO9632007-A1.
XX
XX 17-OCT-1996.
PD
PD
XX 15-APR-1996; 96WO-US05272.
PF
PF
XX 13-APR-1995; 95IL-0113373.
PR
PR
XX (YEDA ) YEDA RES & DEV CO LTD.
PA (YISS ) YISSUM RES & DEV CO.
PA
XX Eshed Y, Fluhr R, Ori N, Paran I, Zamir D;
PI WPI; 1996-476739/47.
XX N-PSDB; AAT42134.
XX
XX New DNA conferring resistance to Fusarium sp. - used for producing
PT disease-resistant tomato plants and for identifying new resistance
PT genes and diseases
PT
XX Claim 1; Fig 4A-C; 50pp; English.
PS
XX This I2C-1 protein is encoded by a sequence from the I2C multigene
CC family from the I2 Fusarium wilt resistance locus of tomato, and
CC confers resistance to Fusarium oxysporum f.sp. lycopersici race-2.
CC This sequence and I2C-2 (AAW03665) are encoded by genes from a locus
CC completely linked to I2, and show structural similarity with other
CC resistance proteins. The protein has a conserved N-terminal
CC nucleotide-binding domain (the P-loop) and 5 other conserved
CC domains of unknown function. At least half the C-terminus is
CC composed of leucine-rich repeats, which may be responsible for
CC specificity of interaction, either with a pathogen protein
CC component, or with downstream factors involved with signal
CC transduction. There does not appear to be a transmembrane domain,
CC indicating an intracellular location. A putative leucine zipper
CC domain has been predicted. I2C genes may be inserted in a cosmid
CC vector for expression in a tomato transgenic plant, to confer
CC disease-resistance, or may be used as restriction fragment length
CC polymorphism probes for screening for and selective breeding of
CC tomato or Solanaceae plants with disease-resistance.
XX
XX Sequence 1220 AA;

Query Match 16.0%; Score 8; DB 17; Length 1220;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 EQLRTLTP 24
Db 554 eqlrtltp 561
|||||

RESULT 4
AAW03665
ID AAW03665 standard; Protein: 1240 AA.
XX
XX AAW03665;
XX
XX 22-FEB-1997 (first entry)
DT
DE I2C-2 protein conferring Fusarium wilt disease-resistance.
XX
XX Tomato; I2C-2; Fusarium wilt; disease-resistance; Fusarium oxysporum;
KW P-loop; leucine-rich repeat; transgenic plant; screening;
KW restriction fragment length polymorphism; crop improvement;
KW Solanaceae.
XX
XX Lycopersicon esculentum.
OS
XX

```

```

FH Key Location/Qualifiers
FT Domain 201..208
FT /note= "P-loop"
FT
FT Domain 273..282
FT /note= "Conserved motif of unknown function"
FT
FT Domain 304..311
FT /note= "Conserved motif of unknown function"
FT
FT Domain 367..372
FT /note= "Conserved motif of unknown function"
FT
FT Misc-difference 387..388
FT /note= "Conserved motif of unknown function"
FT
FT Domain 410..415
FT /note= "Conserved motif of unknown function"
FT
FT Domain 488..497
FT /note= "Conserved motif of unknown function"
FT
XX
XX WO9632007-A1.
PN
PN 17-OCT-1996.
PD
PD
XX 15-APR-1996; 96WO-US05272.
PF
PF
XX 13-APR-1995; 95IL-0113373.
PR
PR
XX (YEDA ) YEDA RES & DEV CO LTD.
PA (YISS ) YISSUM RES & DEV CO.
PA
XX Eshed Y, Fluhr R, Ori N, Paran I, Zamir D;
PI WPI; 1996-476739/47.
XX N-PSDB; AAT42135.
XX
XX New DNA conferring resistance to Fusarium sp. - used for producing
PT disease-resistant tomato plants and for identifying new resistance
PT genes and diseases
PT
XX Claim 1; Fig 4A-C; 50pp; English.
PS
XX This I2C-2 protein is encoded by a sequence from the I2C multigene
CC family from the I2 Fusarium wilt resistance locus of tomato, and
CC confers resistance to Fusarium oxysporum f.sp. lycopersici race-2.
CC This sequence and I2C-1 (AAW03664) are encoded by genes from a locus
CC completely linked to I2, and show structural similarity with other
CC resistance proteins. The protein has a conserved N-terminal
CC nucleotide-binding domain (the P-loop) and 5 other conserved
CC domains of unknown function. At least half the C-terminus is
CC composed of leucine-rich repeats, which may be responsible for
CC specificity of interaction, either with a pathogen protein
CC component, or with downstream factors involved with signal
CC transduction. There does not appear to be a transmembrane domain,
CC indicating an intracellular location. I2C genes may be inserted
CC in a cosmid vector for expression in a tomato transgenic plant, to
CC confer disease-resistance, or may be used as restriction fragment
CC length polymorphism probes for screening for and selective breeding
CC of tomato or Solanaceae plants with disease-resistance.
XX
XX Sequence 1240 AA;

Query Match 16.0%; Score 8; DB 17; Length 1240;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 EQLRTLTP 24
Db 545 eqlrtltp 552
|||||

RESULT 5
AAW25157
ID AAW25157 standard; Protein: 1266 AA.
XX
XX AAW25157;
AC

```

PA (AFAR/) AFAR D E.  
 PA (HUBE/) HUBERT R S.  
 PA (RAIT/) RAITANO A B.  
 XX  
 XX Afar DE, Hubert RS, Raitano AB;  
 XX  
 XX WPI: 2000-237872/20.  
 DR N-PSDB; AA294275.  
 XX  
 XX Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -  
 XX  
 XX Claim 1; Fig 2A-D; 62pp; English.  
 XX  
 XX This sequence is that of human PHELIX, a novel basic Helix Loop  
 CC Helix protein thought to act as a transcription factor. PHELIX  
 CC normally exhibits a testis-specific expression pattern but is  
 CC up-regulated in prostate and other types of cancer. The invention  
 CC provides diagnostic and therapeutic methods useful in the  
 CC management of various cancers which express PHELIX, including  
 CC prostate cancer, bladder cancer, ovarian cancer and testicular  
 CC cancer, including therapies aimed at inhibition the transcription,  
 CC translation, processing or function of PHELIX. The expression  
 CC pattern of PHELIX suggests that is an ideal target for a cancer  
 CC vaccine approach to prostate cancer. PHELIX protein can also be  
 CC used to screen for agonists and antagonists of therapeutic value  
 CC and to raise antibodies.  
 XX  
 XX Sequence 405 AA:

Query Match 100.0%; Score 50; DB 21; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-43;  
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSSKEKLRRERIKYCEQLRTLLPVYKGRKNDAAASVLEATVDYVKYIREK 50  
 DB 140 hsksklrrerikyccqrltlpvykgrkndaasvleatvdvykyirek 189

RESULT 2  
 AAY79270  
 ID AAY79270 standard; Peptide; 15 AA.  
 XX  
 XX AC AAY79270;  
 XX  
 XX 03-JUL-2000 (first entry)  
 XX  
 XX PHELIX peptide used to raise antibody.  
 XX  
 XX PHELIX; human; testis-specific; transcription factor;  
 KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
 KW therapy; diagnosis; vaccine; antibody.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200012709-A2.  
 PN  
 XX  
 XX 09-MAR-2000.  
 PD  
 XX  
 XX 31-AUG-1999; 99WO-US20137.  
 PF  
 XX  
 XX 31-AUG-1998; 98US-0098610.  
 PR  
 XX 31-OCT-1998; 98US-0106524.  
 PR  
 XX (UROC-) UROGENESYS INC.  
 PA (AFAR/) AFAR D E.  
 PA (HUBE/) HUBERT R S.  
 PA (RAIT/) RAITANO A B.  
 XX  
 XX Afar DE, Hubert RS, Raitano AB;  
 PI  
 XX

DR WPI: 2000-237872/20.  
 XX  
 XX Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -  
 XX  
 XX Example 5; Page 36; 62pp; English.  
 XX  
 XX The present sequence is that of a peptide derived from human  
 CC PHELIX (see AAY79269), a novel transcription factor that is  
 CC normally expressed only in testis tissue, but which is up-regulated  
 CC in prostate and some other cancers. The peptide was conjugated to  
 CC keyhole limpet haemocyanin and used to raise polyclonal antiserum  
 CC in rabbit. The antiserum demonstrated specificity for PHELIX and  
 CC may therefore be useful for assessing the expression of PHELIX in  
 CC patient samples.  
 XX  
 XX Sequence 15 AA;

Query Match 28.0%; Score 14; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-08;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSSKEKLRRERIKY 14  
 DB 1 hsksklrreriky 14

RESULT 3  
 AAW03664  
 ID AAW03664 standard; Protein; 1220 AA.  
 XX  
 XX AC AAW03664;  
 XX  
 XX 22-FEB-1997 (first entry)  
 XX  
 XX I2C-1 protein conferring Fusarium wilt disease-resistance.  
 DE  
 XX Tomato; I2C-1; Fusarium wilt; disease-resistance; Fusarium oxysporum;  
 KW P-loop; leucine zipper; leucine-rich repeat; transgenic plant;  
 KW restriction fragment length polymorphism; screening;  
 KW crop improvement; Solanaceae.  
 XX  
 XX Lycopersicon esculentum.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Region 66..71  
 FT /note= "Repeat sequence"  
 FT Region 103..108  
 FT /note= "Repeat sequence"  
 FT Domain 200..207  
 FT /note= "P-loop"  
 FT Region 256..267  
 FT /note= "Repeat sequence"  
 FT Region 269..280  
 FT /note= "Repeat sequence"  
 FT Domain 286..295  
 FT /note= "Conserved motif of unknown function"  
 FT Domain 317..324  
 FT /note= "Conserved motif of unknown function"  
 FT Domain 380..385  
 FT /note= "Conserved motif of unknown function"  
 FT Misc-difference 400..401  
 FT /note= "Conserved motif of unknown function"  
 FT Domain 419..430  
 FT /note= "Conserved motif of unknown function"  
 FT Domain 498..506  
 FT /note= "Conserved motif of unknown function"  
 FT Domain 559..623  
 FT /note= "Leucine-rich repeat region"  
 FT Domain 624..645  
 FT /note= "Putative leucine zipper domain"

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:47:12 ; Search time 65.58 Seconds  
(without alignments)  
46.221 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_189

Perfect score: 50

Sequence: 1 HSSKEKLRRIKCYCCQLR.....NDAAVLEATVDYVKYIREK 50

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_0601.\*

- 1: /cgnl\_9/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /cgnl\_9/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /cgnl\_9/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /cgnl\_9/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /cgnl\_9/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /cgnl\_9/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /cgnl\_9/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /cgnl\_9/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /cgnl\_9/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /cgnl\_9/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /cgnl\_9/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /cgnl\_9/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /cgnl\_9/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /cgnl\_9/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /cgnl\_9/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /cgnl\_9/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /cgnl\_9/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /cgnl\_9/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /cgnl\_9/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /cgnl\_9/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /cgnl\_9/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	405	21	AA19269 Human testis-speci
2	14	28.0	15	21	AA19270 PHELI peptide use
3	8	16.0	1220	17	AAW03664 I2C-1 protein conf
4	8	16.0	1240	17	AAW03665 I2C-2 protein conf
5	8	16.0	1266	18	AAW25157 Tomato immunity 2
6	8	16.0	1266	21	AAW07754 Amino acid sequenc
7	7	14.0	135	21	AAW43791 Arabidopsis thalia
8	7	14.0	170	20	AAW29197 Amino acid sequenc
9	7	14.0	250	21	AAW43790 Arabidopsis thalia
10	7	14.0	291	21	AAW15758 Arabidopsis thalia
11	7	14.0	341	21	AAW15757 Arabidopsis thalia

12	7	14.0	443	21	AAW15756 Arabidopsis thalia
13	7	14.0	477	20	AAW37305 Amino acid sequenc
14	7	14.0	514	13	AAW28401 Bovine lung 59 kd
15	7	14.0	514	16	AAW69715 Cyclic-GMP stimula
16	7	14.0	514	18	AAW18037 Bovine lung 59 kDa
17	7	14.0	514	18	AAW11240 Calcium/calmodulin
18	7	14.0	514	19	AAW77037 Bovine lung Ca2+/c
19	7	14.0	514	19	AAW71221 59 kDa CaM-PDE fro
20	7	14.0	514	19	AAW60749 59 kDa bovine CaM-
21	7	14.0	514	21	AAW80972 Bovine lung 59 kd
22	7	14.0	530	13	AAW28395 Bovine brain Cam P
23	7	14.0	530	16	AAW69711 Cyclic-GMP stimula
24	7	14.0	530	18	AAW18036 Bovine brain 61 kd
25	7	14.0	530	18	AAW11243 61 kd brain calciu
26	7	14.0	530	19	AAW77036 Bovine brain Ca2+/
27	7	14.0	530	19	AAW71220 61 kDa CaM-PDE DNA
28	7	14.0	530	19	AAW60747 61 kDa bovine CaM-
29	7	14.0	535	13	AAW80968 Protein encoded by
30	7	14.0	535	13	AAW28411 Cyclic-GMP stimula
31	7	14.0	535	16	AAW69731 Human brain 61 kDa
32	7	14.0	535	18	AAW18039 Hippocampus calciu
33	7	14.0	535	18	AAW11255 Human 61 kDa CaM-P
34	7	14.0	535	19	AAW71227 Human Ca2+/calmodu
35	7	14.0	535	19	AAW77043 Amino acid sequenc
36	7	14.0	535	19	AAW60755 Human 61 kd CaM-PD
37	7	14.0	535	21	AAW80988 Comamonas testoste
38	6	12.0	27	15	AAW15385 Human secreted pro
39	6	12.0	91	21	AAW03544 Bordetella pertuss
40	6	12.0	125	21	AAW14135 Human leukotriene
41	6	12.0	150	17	AAW02605 Human leukotriene C
42	6	12.0	150	17	AAW90285 Arabidopsis thalia
43	6	12.0	152	21	AAW54945 Sequence of full-l
44	6	12.0	153	9	AAW81879 Sequence of human
45	6	12.0	153	9	AAW80929

ALIGNMENTS

RESULT	1
AAW79269	
ID	AAW79269 standard; Protein; 405 AA.
XX	
AC	AAW79269;
XX	
DT	03-JUL-2000 (first entry)
XX	
DE	Human testis-specific transcription factor PHELI.
XX	
KW	PHELI; human; testis-specific; transcription factor; prostate cancer; bladder cancer; ovary cancer; testicular cancer; therapy; diagnosis; vaccine.
KW	
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	Peptide
FT	Peptide
FT	Domain
FT	Domain
XX	
PN	WO200012709-A2.
XX	
PD	09-MAR-2000.
XX	
PF	31-AUG-1999; 99WO-US20137.
XX	
PR	31-AUG-1998; 98US-0098610.
PR	31-OCT-1998; 98US-0106524.
XX	
PA	(UROC-) UROGENESYS INC.





QY 3 SKEKL 7  
|||||  
Db 19 SKEKL 23

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 SSKEK 6  
|||||  
Db 3 SSKEK 7

## RESULT 14

Q9UJ31 PRELIMINARY; PRT; 47 AA.  
AC Q9UJ31;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE 60S RIBOSOMAL PROTEIN K5 (FRAGMENT).  
GN RPK5+  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-968 H90;  
RA Ding D.;  
RT "Generation and analysis of GFP-gene fusion library of fission yeast."  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB027845; BAA87149.1; -  
DR InterPro; IPR002171; -  
DR Pfam; PF00181; Ribosomal\_L2; 1.  
KW Ribosomal protein.  
FT NON\_TER 47  
SQ SEQUENCE 47 AA; 5348 MW; 1546CB53CAA187BB CRC64;

Query Match 10.0%; Score 5; DB 3; Length 47;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 QLRTL 22  
|||||  
Db 27 QLRTL 31

## RESULT 15

O20181 PRELIMINARY; PRT; 47 AA.  
AC O20181;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE ORF47B.  
OS Chlorella vulgaris.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
OC Chlorellaceae; Chlorella.  
OX NCBI\_TaxID=3077;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97303241; PubMed=9159184;  
RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito S.,  
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,  
RA Inamura A., Yoshinaga K., Sugitara M.;  
RT "Complete nucleotide sequence of the chloroplast genome from the green  
alga Chlorella vulgaris: the existence of genes possibly involved in  
chloroplast division."  
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
DR EMBL; AB001684; BAA57959.1; -  
KW Chloroplast.  
SQ SEQUENCE 47 AA; 5867 MW; B161224E524F3802 CRC64;

Query Match 10.0%; Score 5; DB 8; Length 47;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;

RESULT 10  
Q9KKN8 PRELIMINARY; PRT; 40 AA.  
AC Q9KKN8;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE HYPOTHETICAL PROTEIN VCA1064.  
GN VCA1064.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-EL TOR N16961 / SEROTYPE O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
RT cholerae".  
RL Nature 406:477-483(2000).  
DR EMBL; AE004432; AAF96958.1;  
DR TIGR; VCA1064;  
KW Hypothetical protein.  
SQ SEQUENCE 40 AA; 4465 MW; 9861C2D6CBF96BAF CRC64;

Query Match 10.0%; Score 5; DB 2; Length 40;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 27 KGRKN 31  
Db 35 KGRKN 39

RESULT 11  
Q50817 PRELIMINARY; PRT; 42 AA.  
AC Q50817;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE 65 KDA ANTIGEN (CELL WALL PROTEIN A) GENE (CELL WALL PROTEIN A).  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ERDMAN;  
RX MEDLINE=87137260; PubMed=3029018;  
RA Shinnick T.M.;  
RT "The 65-kilodalton antigen of Mycobacterium tuberculosis";  
RT J. Bacteriol. 169:1080-1088(1987).  
DR EMBL; M15467; AA888239.1;  
SQ SEQUENCE 42 AA; 4437 MW; 6A1521E60DBC0E29 CRC64;

Query Match 10.0%; Score 5; DB 2; Length 42;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVL 37  
Db 11 AASVL 15

RESULT 12  
O77577 PRELIMINARY; PRT; 42 AA.  
AC O77577;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE UROCORTIN PRECURSOR (FRAGMENT).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cepoi D., Sutton S., Vale W.W.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 14-42 FROM N.A.  
RA Balgent S.M., Lowry P.J.;  
RT "The cloning of ovine urocortin";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF051807; AAC37288.1;  
DR EMBL; AF084258; AAC33478.1;  
DR InterPro; IPR000187;  
DR Pfam; PF00473; CRF; 1.  
DR PROSITE; PS00511; CRF; 1.  
DR SMART; SM00039; CRF; 1.  
FT NON\_TER 1  
SQ SEQUENCE 42 AA; 4893 MW; 31D6139D4937C9E2 CRC64;

Query Match 10.0%; Score 5; DB 6; Length 42;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 19 LRTLL 23  
Db 14 LRTLL 18

RESULT 13  
O20145 PRELIMINARY; PRT; 42 AA.  
AC O20145;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE ORF42A.  
OS Chlorella vulgaris.  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
OC Chlorellaceae; Chlorella.  
OX NCBI\_TaxID=3077;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97303241; PubMed=9159184;  
RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,  
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,  
RA Inamura A., Yoshinaga K., Sugura M.;  
RT "Complete nucleotide sequence of the chloroplast genome from the green  
RT alga Chlorella vulgaris: the existence of genes possibly involved in  
RT chloroplast division";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
DR EMBL; AB001684; BAA57895.1;  
KW Chloroplast.  
SQ SEQUENCE 42 AA; 5252 MW; 445B221534CACD71 CRC64;

Query Match 10.0%; Score 5; DB 8; Length 42;  
Best Local Similarity 100.0%; Pred. No. 5.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VLEAT 40  
DB 24 VLEAT 28

## RESULT 6

Q9RCT4 ID Q9RCT4 PRELIMINARY; PRT; 35 AA.  
AC Q9RCT4;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE M-PROTEIN (FRAGMENT).  
GN M.  
OS Streptococcus equi.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1336;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=E;  
RA Chatter N., Talbot N., Newton R., Verheyen K.;  
RT "Independent multiple emergence of Streptococcus equi with truncated  
RT M-like proteins in long-term carriers from different outbreaks of  
RT equine strangles.";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ249874; CAB64611.1; -  
FT NON\_TER 1 1  
FT NON\_TER 35 35  
SQ SEQUENCE 35 AA; 3540 MW; 925BAA2F0C45DC4 CRC64;

Query Match 10.0%; Score 5; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVL 37  
DB 2 AASVL 6

## RESULT 7

Q9R5H9 ID Q9R5H9 PRELIMINARY; PRT; 35 AA.  
AC Q9R5H9;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE SURFACE ARRAY PROTEIN (FRAGMENT).  
OS Aeromonas hydrophila.  
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;  
OC Aeromonas.  
OX NCBI\_TaxID=644;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92407495; PubMed=1382113;  
RA Kokka R.P., Vedros N.A., Janda J.M.;  
RT "Immunochemical analysis and possible biological role of an Aeromonas  
RT hydrophila surface array protein in septicemia.";  
RL J. Gen. Microbiol. 138:1229-1236(1992).  
SQ SEQUENCE 35 AA; 3537 MW; 2E8BDDC978B4C CRC64;

Query Match 10.0%; Score 5; DB 2; Length 35;  
Best Local Similarity 100.0%; Pred. No. 4.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 DAASV 36  
DB 5 DAASV 9

## RESULT 8

Q9RCT8 ID Q9RCT8 PRELIMINARY; PRT; 39 AA.  
AC Q9RCT8;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE M-PROTEIN (FRAGMENT).  
GN M.  
OS Streptococcus equi.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1336;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3;  
RA Chatter N., Talbot N., Newton R., Verheyen K.;  
RT "Independent multiple emergence of Streptococcus equi with truncated  
RT M-like proteins in long-term carriers from different outbreaks of  
RT equine strangles.";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ249870; CAB64607.1; -  
FT NON\_TER 1 1  
FT NON\_TER 39 39  
SQ SEQUENCE 39 AA; 4016 MW; 0C74C9D7AC21AF60 CRC64;

Query Match 10.0%; Score 5; DB 2; Length 39;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVL 37  
DB 2 AASVL 6

## RESULT 9

Q19688 ID Q19688 PRELIMINARY; PRT; 39 AA.  
AC Q19688;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE HLA-B\*27 VARIANT EXON 2 (ALPHA1 DOMAIN) (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Blasczyk R., Weber M., Salama A.;  
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X83727; CAA58698.1; -  
DR HSSP; P01891; 2HLA.  
DR InterPro; IPR001039; -  
DR Pfam; PF00129; MHC\_I; 1.  
KW MHC.  
FT NON\_TER 1 1  
FT NON\_TER 39 39  
SQ SEQUENCE 39 AA; 4748 MW; F5E3FD2A53138908 CRC64;

Query Match 10.0%; Score 5; DB 7; Length 39;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
DB 27 LRTLL 31

Db 17 SSKEKL 22

## RESULT 2

Q9RCT6 PRELIMINARY; PRT; 29 AA.  
 AC Q9RCT6; 29 AA.  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
 DE M-PROTEIN (FRAGMENT)  
 GN M.  
 OS Streptococcus equi.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C;  
 RA Chanter N., Talbot N., Newton R., Verheyen K.;  
 RT "Independent multiple emergence of Streptococcus equi with truncated  
 RT M-like proteins in long-term carriers from different outbreaks of  
 RT equine strangles."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ249872; CAB64609.1;  
 FT NON\_TER 1  
 FT NON\_TER 29  
 FT NON\_TER 29  
 SQ SEQUENCE 29 AA; 2860 MW; EEFBE55184CCLFE7 CRC64;

Query Match 10.0%; Score 5; DB 2; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVL 37  
 |||||  
 Db 2 AASVL 6

## RESULT 3

Q9RCT5 PRELIMINARY; PRT; 31 AA.  
 AC Q9RCT5; 31 AA.  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)  
 DE M-PROTEIN (FRAGMENT)  
 GN M.  
 OS Streptococcus equi.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D;  
 RA Chanter N., Talbot N., Newton R., Verheyen K.;  
 RT "Independent multiple emergence of Streptococcus equi with truncated  
 RT M-like proteins in long-term carriers from different outbreaks of  
 RT equine strangles."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ249873; CAB64610.1;  
 FT NON\_TER 1  
 FT NON\_TER 31  
 FT NON\_TER 31  
 SQ SEQUENCE 31 AA; 3061 MW; C06E2E6898306D01 CRC64;

Query Match 10.0%; Score 5; DB 2; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 AASVL 37  
 |||||  
 Db 2 AASVL 6

## RESULT 4

Q9KP96 PRELIMINARY; PRT; 31 AA.  
 AC Q9KP96; 31 AA.  
 DT 01-OCT-2000 (TRENBLrel. 15, Created)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
 DE HYPOTHETICAL PROTEIN VC2477.  
 GN VC2477.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae."  
 RL Nature 406:477-483(2000).  
 DR EMBL; AE004317; AAF95619.1;  
 DR TIGR; VC2477;  
 KW Hypothetical protein.  
 SQ SEQUENCE 31 AA; 3493 MW; 33DC6391D3FBD4F CRC64;

Query Match 10.0%; Score 5; DB 2; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LRTLL 23  
 |||||  
 Db 16 LRTLL 20

## RESULT 5

Q9W7P0 PRELIMINARY; PRT; 31 AA.  
 AC Q9W7P0; 31 AA.  
 DT 01-NOV-1999 (TRENBLrel. 12, Created)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
 DE CLASS 7A MYOSIN (FRAGMENT)  
 OS Morone saxatilis (Striped bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 OC Moronidae; Morone.  
 OX NCBI\_TaxID=34816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=RPE;  
 RA Hillman D.W., Bost-Usinger L., Cheng J., Burnside B.;  
 RT "Multiple Myosins are Expressed in Fish RPE and Retina."  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF002867; AAD41448.1;  
 DR InterPro: IPR001609;  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Prodom: PD000355; -; 1.  
 FT NON\_TER 1  
 FT NON\_TER 31  
 FT NON\_TER 31  
 SQ SEQUENCE 31 AA; 3465 MW; 594C38532AB69B04 CRC64;

Query Match 10.0%; Score 5; DB 13; Length 31;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:55:36 : Search time 17.57 Seconds  
(without alignments)  
376.508 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_189

Perfect score: 50

Sequence: 1 HSSKEKLRRIKYCCQQLR.....NDAASVLEATVDYVKYIREK 50

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 25325

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_16:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	12.0	39	Q9UBN9	Q9ubn9 homo sapien
2	5	10.0	29	Q9RCT6	Q9rct6 streptococc
3	5	10.0	31	Q9RCT5	Q9rct5 streptococc
4	5	10.0	31	Q9KP96	Q9kp96 vibrio chol
5	5	10.0	31	Q9WP70	Q9wp70 morone saxa
6	5	10.0	35	Q9RCT4	Q9rct4 streptococc
7	5	10.0	35	Q9R5H9	Q9r5h9 aeromonas h
8	5	10.0	39	Q9RCT8	Q9rct8 streptococc
9	5	10.0	39	Q9P688	Q9p688 homo sapien
10	5	10.0	40	Q9KKN8	Q9kkn8 vibrio chol
11	5	10.0	42	Q50817	Q50817 mycobacteri
12	5	10.0	42	Q77577	Q77577 ovis aries
13	5	10.0	42	Q20145	Q20145 chlorella v
14	5	10.0	47	Q9UW31	Q9uw31 schizosacch
15	5	10.0	47	Q20181	Q20181 chlorella v
16	5	10.0	47	Q9QCM2	Q9qcm2 borna disea
17	5	10.0	47	Q9QCL9	Q9qcl9 borna disea
18	5	10.0	47	Q9QCL6	Q9qcl6 borna disea
19	5	10.0	47	Q9QCL3	Q9qcl3 borna disea

20	5	10.0	47	14	Q9QCL0	Q9qcl0 borna disea
21	5	10.0	47	14	Q9QCK7	Q9qck7 borna disea
22	5	10.0	47	14	Q9QCK4	Q9qck4 borna disea
23	5	10.0	47	14	Q9QCK1	Q9qck1 borna disea
24	5	10.0	50	2	Q9FBJ5	Q9fbj5 streptomyce
25	4	8.0	9	11	Q9QWG2	Q9qwg2 mus musculu
26	4	8.0	11	3	Q9HFN8	Q9hfn8 candida rug
27	4	8.0	11	4	Q9UEL0	Q9uel0 homo sapien
28	4	8.0	11	10	Q9S8X4	Q9s8x4 glycine max
29	4	8.0	12	14	Q85631	Q85631 avian retro
30	4	8.0	13	2	Q47693	Q47693 escherichia
31	4	8.0	13	4	Q9UET3	Q9uet3 homo sapien
32	4	8.0	14	11	Q9Z0G5	Q9z0g5 mus musculu
33	4	8.0	14	14	O10229	O10229 human immun
34	4	8.0	14	14	O10230	O10230 human immun
35	4	8.0	14	14	O10235	O10235 human immun
36	4	8.0	15	2	Q9R5T1	Q9r5t1 flavobacter
37	4	8.0	15	5	Q9RTW3	Q9rtw3 crithidia f
38	4	8.0	15	14	Q97090	Q97090 human immun
39	4	8.0	15	14	Q97092	Q97092 human immun
40	4	8.0	15	14	Q97094	Q97094 human immun
41	4	8.0	15	14	Q97098	Q97098 human immun
42	4	8.0	15	14	Q79359	Q79359 human immun
43	4	8.0	16	2	Q48417	Q48417 klebsiella
44	4	8.0	16	4	Q9UBI5	Q9ubi5 homo sapien
45	4	8.0	16	4	Q9UD21	Q9ud21 homo sapien

#### ALIGNMENTS

RESULT 1

Q9UBN9  
ID Q9UBN9 PRELIMINARY; PRT; 39 AA.  
AC Q9UBN9;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE E6-AP UBIQUITIN-PROTEIN LIGASE (FRAGMENT).  
GN UBE3A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE=98126441; PubMed=9465301;  
RA Kishino T., Wagstaff J.;  
RT "Genomic organization of the UBE3A/E6-AP gene and related  
pseudogenes.";  
RL Genomics 47:101-107(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Hennies H.C., Buerger J., Sperling K., Reis A.;  
RT "Mutations in the E6-AP gene (UBE3A) in patients with Angelman  
syndrome.";  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF009341; AAC39580.1; -;  
DR EMBL; AJ001113; CAA04540.1; -;  
DR InterPro; IPR000569; -;  
DR PROSITE; PS50237; HECT; 1.  
KW Ligase.  
FT NON\_TER 1 1  
SQ SEQUENCE 39 AA; 4441 MW; 763722F374FA7193 CRC64;

Query Match 12.0%; Score 6; DB 4; Length 39;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKEKL 7  
|||||

```

RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Felschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervilave A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32844; AAC23364.1; -
DR TIGR: H11717; -
KW Hypothetical protein.
SQ SEQUENCE 32 AA; 3666 MW; 4A117C0B1E768C2 CRC64;

```

```

Query Match          8.0%; Score 4; DB 1; Length 32;
Best Local Similarity 100.0%; Pred.No. 8.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 QLRT 21
   ||||
Db 13 QLRT 16

```

Search completed: September 15, 2001, 12:58:07  
Job time: 106 sec

OC Elapidae; Elapinae; Micrurus.  
 RN NCBI\_TaxID=8635;  
 RC SEQUENCE.  
 RC TISSUE=Venom;  
 RA Mochica-Morales J., Martin B.M., Zamudio F.Z., Possani L.D.;  
 RT "Isolation and characterization of three toxic phospholipases from  
 the venom of the coral snake *Micrurus nigrocinctus*.";  
 RL Toxicon 28:616-617(1990).  
 CC -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE  
 CC 2-ACYL GROUPS IN 3-SN-PHOSPHOLIPIDS.  
 CC -!- FUNCTION: INHIBITS NEUROMUSCULAR TRANSMISSION BY BLOCKING  
 CC ACETYLCHOLINE RELEASE FROM THE NERVE TERMINI. ACT PRESYNAPTICALLY.  
 CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-  
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.  
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.  
 DR PIR; A35948; A35948.  
 DR HSP; P00598; IPOB.  
 DR InterPro; IPR001211; .  
 DR Pfam; PF00068; phoslip; 1.  
 DR PROSITE; PS00118; PA2\_HIS; PARTIAL.  
 DR PROSITE; PS00119; PA2\_ASP; PARTIAL.  
 KW Hydrolase; Lipid degradation; Calcium; Presynaptic neurotoxin; Venom;  
 KW Multigene family.  
 FT NON\_TER 27 27  
 SQ SEQUENCE 27 AA; 3314 MW; 38637EAC600F49A0 CRC64;

Query Match 8.0%; Score 4; DB 1; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 SVLE 38  
 DB 16 SVLE 19

RESULT 13  
 IDH COMTE  
 ID DIDH\_COMTE STANDARD; PRT; 30 AA.  
 AC P80702;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE 3-ALPHA-HYDROXYSTEROID DEHYDROGENASE (EC 1.1.1.50) (3-ALPHA-HSD)  
 OS (HYDROXYPROSTAGLANDIN DEHYDROGENASE) (HSD28) (FRAGMENT).  
 DE Comamonas testosteroni (pseudomonas testosteroni).  
 OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.  
 RN NCBI\_TaxID=285;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=ATCC 11996;  
 RX MEDLINE=97100200; PubMed=8944761;  
 RA Oppermann U.C.T., Maser E.;  
 RT "Characterization of a 3 alpha-hydroxysteroid dehydrogenase/carbonyl  
 reductase from the gram-negative bacterium *Comamonas testosteroni*.";  
 RL Eur. J. Biochem. 241:744-749(1996).  
 CC -!- FUNCTION: ALONG WITH THE 3 ALPHA-HYDROXYSTEROID DEHYDROGENASE AND  
 CC 3-OXO-REDUCTASE ACTIVITIES TOWARDS A VARIETY OF CIS OR TRANS FUSED  
 CC A/B RING STEROIDS, IT ALSO REDUCES SEVERAL XENOBIOTIC CARBONYL  
 CC COMPOUNDS, INCLUDING A METRAPONE-BASED CLASS OF INSECTICIDES, TO  
 CC THE RESPECTIVE ALCOHOL METABOLITES.  
 CC -!- CATALYTIC ACTIVITY: ANDROSTERONE + NAD(P)(+) -  
 CC 5-ALPHA-ANDROSTANE-3,17-DIONE + NAD(P)H.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- INDUCTION: BY STEROIDS.  
 CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.  
 DR InterPro; IPR002198; .  
 DR PROSITE; PS00061; ADH\_SHORT; PARTIAL.  
 KW Oxidoreductase; NAD.  
 FT DOMAIN 10 >30 INVOLVED IN COFACTOR BINDING  
 (BY SIMILARITY).

FT NON\_TER 30 30  
 SQ SEQUENCE 30 AA; 2829 MW; 065E9CF03F1C5A29 CRC64;

Query Match 8.0%; Score 4; DB 1; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VLEA 39  
 DB 19 VLEA 22

RESULT 14  
 PSAM\_ODOSI  
 ID PSAM\_ODOSI STANDARD; PRT; 30 AA.  
 AC P49487;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE PHOTOSYSTEM I REACTION CENTRE SUBUNIT XII (PSI-M).  
 PSAM.  
 GN Odontella sinensis.  
 OS Chloroplast.  
 OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;  
 OC Biddulphiophycidae; Eupodiscales; Eupodiscaeae; Odontella.  
 OX NCBI\_TaxID=2839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;  
 RT "The chloroplast genome of a chlorophyll a+c-containing alga,  
*Odontella sinensis*.";  
 RL Plant Mol. Biol. Rep. 13:336-342(1995).  
 CC -!- SIMILARITY: BELONGS TO THE PSAM FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; 267753; CAA91676.1; .  
 DR Mendel; 2595; ODOSI; psam; 1.  
 KW Photosystem I; Photosynthesis; Chloroplast.  
 SQ SEQUENCE 30 AA; 3329 MW; 73FDEB91E4BF634F CRC64;

Query Match 8.0%; Score 4; DB 1; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ASVL 37  
 DB 17 ASVL 20

RESULT 15  
 YH17\_HAEIN  
 ID YH17\_HAEIN STANDARD; PRT; 32 AA.  
 AC P44295;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE HYPOTHETICAL PROTEIN H11717.  
 GN H11717.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.



CC RHYTHMIC CONTRACTION OF DUODENAL AND COLONIC SMOOTH MUSCLE.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- SIMILARITY: BELONGS TO THE MOTILIN FAMILY.  
DR PIR; S00189; S00189.  
DR PIR; A60313; A60313.  
FW Hormone.  
KT UNSURE  
SQ SEQUENCE 22 AA; 2685 MW; 4BECB840ABE0639F CRC64;

Query Match 8.0%; Score 4; DB 1; Length 22;  
Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 IREK 50  
Db 13 IREK 16

RESULT 9  
FLAL SULSH STANDARD; PRT; 23 AA.  
AC Q9UG66;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE 31/33 KDA FLAGELLIN (FRAGMENT).  
OS Sulfolobus shibatae.  
OC Archaea: Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
OX NCBI\_TaxID=2286;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=B12;  
RX MEDLINE=96146545; PubMed=8550530;  
RA Faguy D.M., Bayley D.P., Kostyukova A.S., Thomas N.A., Jarrell K.F.;  
RT "Isolation and characterization of flagella and flagellin proteins  
from the Thermocacidophilic archaea Thermoplasma volcanium and  
Sulfolobus shibatae";  
RT Sulfolobus shibatae.  
RL J. Bacteriol. 178:902-905(1996).  
CC -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO  
CC FORM THE FILAMENTS OF FLAGELLA.  
CC -!- PTM: GLYCOSYLATED.  
CC -!- SIMILARITY: BELONGS TO THE ARCHAEAL FLAGELLIN FAMILY.  
KW Flagella; Glycoprotein.  
FT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2362 MW; BB372157B44DB3AC CRC64;

Query Match 8.0%; Score 4; DB 1; Length 23;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ASVL 37  
Db 17 ASVL 20

RESULT 10  
CT31\_LITCI STANDARD; PRT; 24 AA.  
AC P81851; P81852; P81853;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CITROPIN 3.1.2 [CONTAINS: CITROPIN 3.1.1; CITROPIN 3.1].  
OS Litoria citropa (Australian blue mountains tree frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;  
OC Litoria.  
OX NCBI\_TaxID=94770;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Skin;

RX MEDLINE=99435977; PubMed=10504394;  
RA Wegener K.L., Wabnitz P.A., Carver J.A., Bowie J.H., Chia B.C.S.,  
RA Wallace J.C., Tyler M.J.;  
RT "Host defence peptides from the skin glands of the Australian blue  
RT mountain tree-frog Litoria citropa. Solution structure of the  
RT antibacterial peptide citropin 1.1";  
RL Eur. J. Biochem. 265:627-637(1999).  
CC -!- TISSUE SPECIFICITY: DORSAL AND SUBMENTAL SKIN GLANDS.  
FT PEPTIDE 1 24 CITROPIN 3.1.2.  
FT PEPTIDE 1 23 CITROPIN 3.1.1.  
FT PEPTIDE 1 22 CITROPIN 3.1.  
SQ SEQUENCE 24 AA; 2614 MW; C9001E295BD0E15D CRC64;

Query Match 8.0%; Score 4; DB 1; Length 24;  
Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KEKL 7  
Db 7 KEKL 10

RESULT 11  
FEDG\_AMEYE STANDARD; PRT; 24 AA.  
ID FEDG\_AMEYE  
AC P80707;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE FORMATE ESTER DEHYDROGENASE, GAMMA CHAIN (EC 1.2.99.-) (FEDH).  
DE (FRAGMENT).  
OS Amycolatopsis methanolica.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;  
OC Amycolatopsis.  
OX NCBI\_TaxID=1814;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=NCIB 11946;  
RX MEDLINE=96140591; PubMed=8554333;  
RA Kim S.W., Luykx D.M.A.M., de Vries S., Duine J.A.;  
RT "A second molybdo-protein aldehyde dehydrogenase from Amycolatopsis  
methanolica NCIB 11946";  
RL Arch. Biochem. Biophys. 325:1-7(1996).  
CC -!- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA AND A GAMMA  
CC CHAIN.  
KW Oxidoreductase.  
FT NON\_TER 24 24  
SQ SEQUENCE 24 AA; 2746 MW; A93A8EA007D0FC6B CRC64;

Query Match 8.0%; Score 4; DB 1; Length 24;  
Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 RTLL 23  
Db 19 RTLL 22

RESULT 12  
PA21\_MICNI STANDARD; PRT; 27 AA.  
ID PA21\_MICNI  
AC P21790;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE PHOSPHOLIPASE A2 ISOZYME 1 (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE  
DE 2-ACYLHYDROLASE) (FRAGMENT).  
OS Micrurus nigrocinctus (Central American coral snake) (Gargantilla).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;

DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE H(2)-FORMING N5,N10-METHYLENETETRAHYDROMETHANOPTERIN DEHYDROGENASE  
 DE (EC 1.12.99.-) (H2-DEPENDENT METHYLENE-HAMPT DEHYDROGENASE)  
 DE (FRAGMENT).  
 GN HMD.  
 OS Methanobacterium wolfei.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanothermobacter.  
 OX NCBI\_TaxID=145261;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92394151; PubMed=1521540;  
 RA Zingibl C., van Dongen W., Schwoerer B., von Buehau R.,  
 RA Richter M., Klein A., Thauer R.K.;  
 RT "H2-forming methylenetetrahydromethanopterin dehydrogenase, a novel  
 RT type of hydrogenase without iron-sulfur clusters in methanogenic  
 RT archaea.";  
 RL Eur. J. Biochem. 208:511-520(1992).  
 CC -!- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROMETHANOPTERIN +  
 CC H(+) = 5,10-METHENYL-TETRAHYDROMETHANOPTERIN + H(2).  
 CC -!- COFACTOR: ZINC (POSSIBLE).  
 CC -!- PATHWAY: INVOLVED IN METHANOGENESIS.  
 CC -!- SUBUNIT: HOMODIMER.  
 KW Oxidoreductase; Methanogenesis; Zinc.  
 FT NON\_TER 1  
 SQ SEQUENCE 19 AA; 1911 MW; 0C17E9D7BF1F97C9 CRC64;  
  
 Query Match 8.0%; Score 4; DB 1; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 32 DAAS 35  
 DB 14 DAAS 17  
  
 RESULT 6  
 YPRB\_SERMA  
 ID YPRB\_SERMA STANDARD; PRT; 20 AA.  
 AC P22581;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE HYPOTHETICAL PROTEIN IN PROB 5'REGION (FRAGMENT).  
 OS Serratia marcescens.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Serratia.  
 OX NCBI\_TaxID=615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SR41;  
 RX MEDLINE=91237315; PubMed=1851803;  
 RA Omori K., Suzuki S., Imai Y., Komatsubara S.;  
 RT "Analysis of the Serratia marcescens probA operon and feedback  
 RT control of proline biosynthesis".  
 RL J. Gen. Microbiol. 137:509-517(1991).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D90351; BAA14363.1; -;  
 DR EMBL; X53086; CAA37253.1; -;  
 DR PIR; S11643; S11643.  
 DR PIR; C49753; C49753.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 20 AA; 2248 MW; 4DD7777735276674 CRC64;

Query Match 8.0%; Score 4; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 4 KEKL 7  
 DB 12 KEKL 15  
  
 RESULT 7  
 CR34\_LITCE  
 ID CR34\_LITCE STANDARD; PRT; 22 AA.  
 AC P56241;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CAERIN 3.4.  
 OS Litoria caerulea.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;  
 OC Litoria.  
 OX NCBI\_TaxID=30344;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Parotoid gland;  
 RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;  
 RT "Peptides from Australian frogs. The structures of the caerins from  
 RT Litoria caerulea".  
 RL J. Chem. Res. 138:910-936(1993).  
 CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL  
 CC CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN  
 CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.  
 CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL  
 CC GLANDS.  
 CC -!- MASS SPECTROMETRY: MW=2452; METHOD=FAB.  
 KW Antibiotic; Amphibian skin; Amidation.  
 FT MOD\_RES 22  
 SQ SEQUENCE 22 AA; 2455 MW; 3AB40B2200D43663 CRC64;  
  
 Query Match 8.0%; Score 4; DB 1; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 47 IREK 50  
 DB 6 IREK 9  
  
 RESULT 8  
 MOTI\_CANFA  
 ID MOTI\_CANFA STANDARD; PRT; 22 AA.  
 AC P19863;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE MOTILIN.  
 GN MLN.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Intestine;  
 RX MEDLINE=83195948; PubMed=6844563;  
 RA Poltras P., Reeve J.R. Jr., Hunkapiller M.W., Hood L.E., Walsh J.H.;  
 RT "Purification and characterization of canine intestinal motilin.";  
 RL Regul. Pept. 5:197-208(1983).  
 CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE REGULATION OF  
 CC INTERDIGESTIVE GASTROINTESTINAL MOTILITY AND INDIRECTLY CAUSES

Db 15 RERIK 19

Query Match 8.0%; Score 4; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UC15\_MAIZE STANDARD; PRT; 14 AA.  
AC P80621;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 245)  
DE (FRAGMENT).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;  
OC Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Coleoptile;  
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
Pernollet J.-C., Zivy M., de Vienne D.;  
RT "The maize two dimensional gel protein database: towards an integrated  
genome analysis program";  
RL Theor. Appl. Genet. 93:997-1005(1996).  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
PROTEIN IS: 4.8, ITS MW IS: 35.7 KDA.  
DR Maize-2DPAGE; P80621; COLEOPTILE.  
DR MaizedB; 123947; -.  
FT NON\_TER 1 1  
FT NON\_TER 14 14  
SQ SEQUENCE 14 AA; 1396 MW; C68949275F404CD2 CRC64;

Query Match 8.0%; Score 4; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 TLIP 24  
Db 3 TLIP 6

UC27\_MAIZE STANDARD; PRT; 15 AA.  
AC P80633;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 688)  
DE (FRAGMENT).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;  
OC Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Coleoptile;  
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,  
Pernollet J.-C., Zivy M., de Vienne D.;  
RT "The maize two dimensional gel protein database: towards an integrated  
genome analysis program";  
RL Theor. Appl. Genet. 93:997-1005(1996).  
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
PROTEIN IS: 6.4, ITS MW IS: 48.4 KDA.  
CC -!- SIMILARITY: TO XENOPUS HISTONE-BINDING PROTEIN N1/N2 AND RABBIT  
AND HUMAN NUCLEAR AUTOANTIGENIC SPERM PROTEIN.  
DR Maize-2DPAGE; P80633; COLEOPTILE.  
DR MaizedB; 123958; -.  
FT NON\_TER 1 1  
FT NON\_TER 15 15

Query Match 8.0%; Score 4; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 YCCE 17  
Db 4 YCCE 7

RESULT 5  
HMD\_METWO STANDARD; PRT; 19 AA.  
ID HMD\_METWO  
AC P32441.  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)

SQ SEQUENCE 15 AA; 1853 MW; CA0E12A5DAED8DC7 CRC64;

Query Match 8.0%; Score 4; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 EQLR 20  
Db 3 EQLR 6

RESULT 4  
HSTB\_ECOLI STANDARD; PRT; 18 AA.  
AC P01560;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE HEAT-STABLE ENTEROTOXIN ST-2 (ST-B).  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=18D / SEROTYPE 0.42:K86:H37;  
RX MEDLINE=81264141; PubMed=7021541;  
RA Chan S.-K., Giannella R.A.;  
RT "Amino acid sequence of heat-stable enterotoxin produced by  
Escherichia coli pathogenic for man";  
RL J. Biol. Chem. 256:7744-7746(1981).  
RN [2]  
RP DISULFIDE BONDS.  
RX MEDLINE=87191003; PubMed=3552731;  
RA Shimonishi Y., Hidaka Y., Koizumi M., Hane M., Aimoto S., Takeda T.,  
Miwatani T., Takeda Y.;  
RT "Mode of disulfide bond formation of a heat-stable enterotoxin (STh)  
produced by a human strain of enterotoxigenic Escherichia coli.";  
RL FEBS Lett. 215:165-170(1987).  
CC -!- FUNCTION: TOXIN WHICH ACTIVATES THE PARTICULATE FORM OF GUANYLATE  
CYCLASE AND INCREASES CYCLIC GMP LEVELS WITHIN THE HOST  
INTESTINAL EPITHELIAL CELLS.  
CC -!- DISEASE: BOTH HEAT-STABLE AND HEAT-LABILE ENTEROTOXINS ARE  
PRODUCED BY PATHOGENIC STRAINS OF E.COLI AND AFFECT THE DIGESTIVE  
TRACT OF MAMMALS.  
CC -!- SIMILARITY: BELONGS TO THE HEAT-STABLE ENTEROTOXIN FAMILY.  
DR PIR; A01823; QHEC2.  
DR HSSP; P01559; IETN.  
DR InterPro; IPR001489; -.  
DR PROSITE; PS00273; ENTEROTOXIN\_H\_STABLE; 1.  
KW Toxin; Enterotoxin.  
FT DISULFID 5 10  
FT DISULFID 6 14  
FT DISULFID 9 17  
SQ SEQUENCE 18 AA; 1978 MW; D0C975F49D500650 CRC64;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:56:21 : Search time 9.52 seconds  
(without alignments)  
179.913 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_189

Perfect score: 50

Sequence: 1 HSSKEKLRRERIKYCCQLR.....NDAASVLEATVDYVKYIREK 50

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3421

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5	10.0	26	1 PUTA_KLEPN	P23725 klebsiella
2	4	8.0	14	1 UC15_MAIZE	P80621 zea mays (m
3	4	8.0	15	1 UC27_MAIZE	P80633 zea mays (m
4	4	8.0	18	1 HSTB_ECOLI	P01560 escherichia
5	4	8.0	19	1 HMD_METRO	P32441 methanobact
6	4	8.0	20	1 YPR6_SERMA	P25581 serratia ma
7	4	8.0	22	1 CR34_LITCE	P56241 littoria cae
8	4	8.0	22	1 MOTI_CANFA	P19863 canis fami
9	4	8.0	23	1 FLA1_SULSH	P19866 sulfolobus
10	4	8.0	24	1 CT31_LITCI	P81851 littoria cit
11	4	8.0	24	1 FEDG_AMEYE	P80707 amycolatops
12	4	8.0	27	1 PA21_MICNI	P21790 micrurus ni
13	4	8.0	30	1 DIDH_COMTE	P80702 comamonas t
14	4	8.0	30	1 PSAM_ODOSI	P49487 odontella s
15	4	8.0	32	1 YH17_HAEIN	P44295 haemophilus
16	4	8.0	34	1 LEC2_CYTSE	P22971 cytisus ses
17	4	8.0	34	1 Y870_HAEIN	P44065 haemophilus
18	4	8.0	35	1 KPRP_PINS	P81664 pinus pinas
19	4	8.0	35	1 LEC1_CYTSE	P22970 cytisus ses
20	4	8.0	35	1 RL7_BUCAP	P41188 buchnera ap
21	4	8.0	36	1 AMPL_PIG	P28839 sus scrofa
22	4	8.0	36	1 RL7_CXBU	O87902 coxiella bu
23	4	8.0	37	1 PSBM_PINTH	P41608 pinus thunb
24	4	8.0	37	1 PYI_CHICK	P25203 gallus gall
25	4	8.0	37	1 RK36_PEA	P07815 pisum sativ
26	4	8.0	37	1 TXOF_HADVE	P81599 hadronyche
27	4	8.0	42	1 GBG7_MOUSE	O61016 mus musculu
28	4	8.0	42	1 V11_BPT7	P03779 bacterioph
29	4	8.0	43	1 AJA4_HORSE	P38031 equus cabal
30	4	8.0	43	1 PSBN_ZAMFU	O9msr1 zamia furtu
31	4	8.0	43	1 TBYI_HUMAN	O14604 homo sapien
32	4	8.0	44	1 RL34_BACST	P23376 bacillus st
33	4	8.0	44	1 RL34_BACST	P05647 bacillus su

34 4 8.0 44 1 RL34\_CXBU P45647 coxiella bu  
35 4 8.0 44 1 RL34\_HAEIN P44370 haemophilus  
36 4 8.0 44 1 RL34\_HELPY P56056 helicobacte  
37 4 8.0 44 1 RL34\_PSEAE P29436 pseudomonas  
38 4 8.0 44 1 RL34\_PSEPU P16498 streptococ  
39 4 8.0 45 1 REPA\_STRPN P13920 streptococ  
40 4 8.0 45 1 RL34\_BACHD Q9rcs3 bacillus ha  
41 4 8.0 46 1 RL34\_CYPAPA P48130 cyanophora  
42 4 8.0 46 1 RK34\_PORPU P51190 porphyra pu  
43 4 8.0 47 1 RL34\_AQUAE O6563 aquifex aeo  
44 4 8.0 47 1 VG60\_BPMLS Q05273 mycobacteri  
45 4 8.0 48 1 RK34\_ODOSI P49566 odontella s

#### ALIGNMENTS

RESULT 1  
ID PUTA\_KLEPN STANDARD; PRT; 26 AA.  
AC P23725;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE BIFUNCTIONAL PUTA PROTEIN [INCLUDES: PROLINE DEHYDROGENASE  
DE (EC 1.5.99.8) (PROLINE OXIDASE); DELTA-1-PYROLINE-5-CARBOXYLATE  
DE DEHYDROGENASE (EC 1.5.1.12) (PSC DEHYDROGENASE)] (FRAGMENT).  
GN PUTA.  
OS Klebsiella pneumoniae.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Klebsiella.  
OX NCBI\_TaxID=573;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91100369; PubMed=1987164;  
RA Chen L.M., Maloy S.;  
RT "Regulation of proline utilization in enteric bacteria: cloning and  
RT characterization of the Klebsiella put control region.";  
RL J. Bacteriol. 173:783-790(1991).  
CC -!- FUNCTION: OXIDIZES PROLINE TO GLUTAMATE FOR USE AS A CARBON AND  
CC NITROGEN SOURCE AND ALSO FUNCTION AS A TRANSCRIPTIONAL REPRESSOR  
CC OF THE PUT OPERON.  
CC -!- CATALYTIC ACTIVITY: L-PROLINE + ACCEPTOR + H(2)O - (S)-1-  
CC PYROLINE-5-CARBOXYLATE + REDUCED ACCEPTOR.  
CC -!- COFACTOR: FAD.  
CC -!- PATHWAY: PROLINE UTILIZATION.  
CC -!- INDUCTION: BY PROLINE, AUTOREPRESSION AND CATABOLITE REPRESSION,  
CC AND IS POTENTIALLY NITROGEN CONTROLLED.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
CC -----  
CC EMBL; M63160; AAA25139.1; -;  
CC InterPro; IPR002086; -;  
CC PROSITE; PS00070; ALDEHYDE\_DEHYDR\_CYS; PARTIAL.  
CC PROSITE; PS00687; ALDEHYDE\_DEHYDR\_GLU; PARTIAL.  
CC Multifunctional enzyme; Oxidoreductase; Flavoprotein; FAD; NAD;  
KW Transcription regulation; Repressor; DNA-binding; Proline metabolism.  
FT NON\_TER 26 26  
SQ SEQUENCE 26 AA; 2824 MW; BB332D0DE504CE19 CRC64;

Query Match 10.0%; Score 5; DB 1; Length 26;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 9 RERIK 13  
|||||

**THIS PAGE BLANK (USPTO)**



US-09-422-869-15/c  
; Sequence 15, Application US/09422869  
; Patent No. 6235481  
; GENERAL INFORMATION:  
; APPLICANT: POLONSKY, KENNETH S.  
; APPLICANT: HORIKAWA, YUKIO  
; APPLICANT: ODA, NAOHISA  
; APPLICANT: COX, NANCY J.  
; APPLICANT: SEENAN, SEAMUS  
; APPLICANT: ZHOU, YUN-PING  
; APPLICANT: OTANI, KENICHI  
; APPLICANT: HANIS, CRAIG L.  
; APPLICANT: BELL, GRAEME I.  
; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES  
; FILE REFERENCE: ARCD:307  
; CURRENT APPLICATION NUMBER: US/09/422,869  
; CURRENT FILING DATE: 1999-10-21  
; EARLIER APPLICATION NUMBER: 60/134,175  
; EARLIER FILING DATE: 1999-05-13  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 1267  
; TYPE: DNA  
; ORGANISM: Human  
US-09-422-869-15

Query Match 1.3%; Score 16; DB 4; Length 1267;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 438 gaagggaagaatcaa 453  
|||||

Db 1201 GAAGGGAAGAATCAA 1186

RESULT 13  
US-08-481-658B-44/c  
; Sequence 44, Application US/08481658B  
; Patent No. 5955075  
; GENERAL INFORMATION:  
; APPLICANT: Zavada, Jan  
; APPLICANT: Pastorekova, Silvia  
; APPLICANT: Pastorek, Jaromir  
; TITLE OF INVENTION: MN Gene and Protein  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leona L. Lauder  
; STREET: 6 Mariposa Court  
; CITY: Tiburon  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94920  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/481,658B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/260,190  
; FILING DATE: 15-JUN-1994  
; NAME: Lauder, Leona L.  
; REGISTRATION NUMBER: 30,863  
; REFERENCE/DOCKET NUMBER: D-0021.3E  
; TELEPHONE: 415-435-2034  
; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1334 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; DESCRIPTION: 6th MN intron  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-481-658B-44

Query Match 1.3%; Score 16; DB 2; Length 1334;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 cttcttcaactaagtgc 33  
|||||

Db 1221 CTTCTTCACTAAGTGC 1206

RESULT 14  
US-08-477-504A-44/c  
; Sequence 44, Application US/08477504A  
; Patent No. 5972353  
; GENERAL INFORMATION:  
; APPLICANT: Zavada, Jan  
; APPLICANT: Pastorekova, Silvia  
; APPLICANT: Pastorek, Jaromir  
; TITLE OF INVENTION: MN Gene and Protein  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leona L. Lauder  
; STREET: 6 Mariposa Court  
; CITY: Tiburon  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94920  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,504A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/260,190  
; FILING DATE: 15-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lauder, Leona L.  
; REGISTRATION NUMBER: 30,863  
; REFERENCE/DOCKET NUMBER: D-0021.3D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-435-2034  
; TELEFAX: 415-435-0727  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1334 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; DESCRIPTION: 6th MN intron  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-477-504A-44

Query Match 1.3%; Score 16; DB 2; Length 1334;  
Best Local Similarity 100.0%; Pred. No. 1e+02;

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,475A
; FILING DATE: December 7, 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, Paul E. Jr.
; REGISTRATION NUMBER: 32011
; REFERENCE/DOCKET NUMBER: PEW/3122/204351
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 935 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE: Lycopersicon chilense
US-08-162-475A-3
```

```

Query Match 1.3%; Score 16; DB 1; Length 935;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 314 actgatctattgcct 329
|||||
Db 331 ACTGATCTATTGCCT 316
```

```

RESULT 10
US-08-162-475A-1/c
; Sequence 1, Application US/08162475A
; Patent No. 5656474
; GENERAL INFORMATION:
; APPLICANT: Zohreh Tabaeizadeh
; TITLE OF INVENTION: A novel endochitinase gene
; TITLE OF INVENTION: Induced by osmotic stress and abscisic acid isolated
; TITLE OF INVENTION: from the wild tomato Lycopersicon chilense Dun.
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARY & CUSHMAN
; STREET: 1100 New York Avenue, N.W., 9th Floor
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 720 kb diskette
; COMPUTER: IBM PS/2, Model 30
; OPERATING SYSTEM: PC-DOS 3.30
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,475A
; FILING DATE: December 7, 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, Paul E. Jr.
; REGISTRATION NUMBER: 32011
; REFERENCE/DOCKET NUMBER: PEW/3122/204351
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE: Lycopersicon chilense
US-08-162-475A-1
```

```

Query Match 1.3%; Score 16; DB 1; Length 966;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 314 actgatctattgcct 329
|||||
Db 360 ACTGATCTATTGCCT 345
```

```

RESULT 11
US-08-882-501-31/c
; Sequence 31, Application US/08882501
; Patent No. 6054269
; GENERAL INFORMATION:
; APPLICANT: GARNIER, Fabien
; APPLICANT: GERBAUD, Guy
; APPLICANT: GALIMAND, Marc
; APPLICANT: COURVALLIN, Patrice
; APPLICANT: DUKTA-MALEN, Sylvie
; APPLICANT: CHARLES, Murielle
; APPLICANT: EVERS, Stefan
; APPLICANT: CASADEWALL, Barbara
; TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR
; TITLE OF INVENTION: DETECTING ENTEROCOCCI AND STREPTOCOCCI BACTERIAL STRAINS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,501
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, Leslie A.
; REGISTRATION NUMBER: 34,872
; REFERENCE/DOCKET NUMBER: 03495.0155-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1095 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus hirae
US-08-882-501-31
```

```

Query Match 1.3%; Score 16; DB 3; Length 1095;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 231 ttaagactgaaacgc 246
|||||
Db 90 TTAAGACTGAAACGC 75
```

RESULT 12



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/760,534A  
FILING DATE: 02-DEC-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/206,186  
FILING DATE: 07-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FI93/00514  
FILING DATE: 01-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: CIMBALA, MICHELE A.  
REGISTRATION NUMBER: 33,851  
REFERENCE/DOCKET NUMBER: 1708.0050004/MAC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26700 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(4378..4443, 22026..22106, 23001..23483,  
LOCATION: 23905..24039, 24251..24418)  
US-08-760-534A-1

Query Match 1.4%; Score 17; DB 3; Length 26700;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1145 gaacaagaatgatttt 1161  
|||||  
DB 21030 GAACAAGAATGATTTT 21014

RESULT 7  
US-08-826-532-15/c  
Sequence 15, Application US/08826532B  
Patent No. 6027923  
GENERAL INFORMATION:  
APPLICANT: Wallace, Robert B.  
TITLE OF INVENTION: Linked Linear Amplification of Nucleic Acids  
FILE REFERENCE: 3239-102P  
CURRENT APPLICATION NUMBER: US/08/826,532B  
CURRENT FILING DATE: 1997-04-02  
EARLIER APPLICATION NUMBER: US 08/475,605  
EARLIER FILING DATE: 1995-06-07  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 15  
LENGTH: 422  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-08-826-532-15

Query Match 1.3%; Score 16; DB 3; Length 422;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 908 aaagtaccccttct 923  
|||||  
DB 146 AAAGTACCCCTTCT 131

RESULT 8  
US-08-896-164-47/c  
Sequence 47, Application US/08896164  
Patent No. 6218521  
GENERAL INFORMATION:  
APPLICANT: OBATA, Yuichi  
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED  
TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR  
TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER  
NUMBER OF SEQUENCES: 87  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/896,164  
FILING DATE: July 17, 1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6218521man D. Hanson  
REGISTRATION NUMBER: 30,945  
REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 822 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-896-164-47

Query Match 1.3%; Score 16; DB 4; Length 822;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 ctgaagaacttgatt 270  
|||||  
DB 592 CTGAAGAAGTGGATT 577

RESULT 9  
US-08-162-475A-3/c  
Sequence 3, Application US/08162475A  
Patent No. 5656474  
GENERAL INFORMATION:  
APPLICANT: Zohreh Tabaeizadeh  
TITLE OF INVENTION: A novel endochitinase gene  
TITLE OF INVENTION: induced by osmotic stress and abscisic acid isolated  
TITLE OF INVENTION: from the wild tomato Lycopersicon chilense Dun.  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: 1100 New York Avenue, N.W., 9th Floor  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 720 kb diskette  
COMPUTER: IBM PS/2, Model 30  
OPERATING SYSTEM: PC-DOS 3.30  
SOFTWARE: WordPerfect 5.1

```
;
; TITLE OF INVENTION: Syndecan Stimulation of Cellular
; TITLE OF INVENTION: Differentiation
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.217
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/206.186
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988.427
; FILING DATE: 01-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1102.0050003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(4378..4443, 22026..22106, 23001..23483,
; LOCATION: 23905..24039, 24251..24418)
; US-08-472-217-1

Query Match 1.4%; Score 17; DB 1; Length 26700;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1145 gaacaaagaatgatttt 1161
|||||
Db 21030 GAACAAAGAATGATTTT 21014

RESULT 5
US-08-488-199-5/c
; Sequence 5, Application US/08488199
; Patent No. 5851993
; GENERAL INFORMATION:
; APPLICANT: Jalkanen, Markku
; APPLICANT: Mali, Markku
; TITLE OF INVENTION: Suppression of Tumor Cell Growth By
; TITLE OF INVENTION: Syndecan-1 Ectodomain
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 New York Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
```

```
;
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488.199
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258.862
; FILING DATE: 13-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1102.0130001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4378..4443
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 22026..22107
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 23002..23483
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 23905..24040
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 24252..24418
; US-08-488-199-5

Query Match 1.4%; Score 17; DB 2; Length 26700;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1145 gaacaaagaatgatttt 1161
|||||
Db 21030 GAACAAAGAATGATTTT 21014

RESULT 6
US-08-760-534A-1/c
; Sequence 1, Application US/08760534A
; Patent No. 6017727
; GENERAL INFORMATION:
; APPLICANT: JALKANEN, MARKKU
; APPLICANT: JAANKOLA, PANU
; APPLICANT: VIHINEN, TAPANI
; TITLE OF INVENTION: SYNDECAN ENHANCER ELEMENT AND SYNDECAN
; TITLE OF INVENTION: STIMULATION OF CELLULAR DIFFERENTIATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

US-08-529-654-3

Query Match 1.5%; Score 18; DB 1; Length 3442;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1074 ttaaaagcacatgttga 1091  
|||||  
Db 3243 ttaaaagcacatgttga 3226

RESULT 2

US-08-913-842-6  
; Sequence 6, Application US/08913842  
; Patent No. 6028250  
; GENERAL INFORMATION:

; APPLICANT: OHBA, Toshiharu  
; APPLICANT: TAKAHASHI, Shuichi  
; APPLICANT: ANMA, Yoshiko  
; APPLICANT: ASADA, Kiyozo  
; APPLICANT: KATO, Ikunoshin  
; TITLE OF INVENTION: PLANT PROMOTER AND METHOD FOR GENE  
; TITLE OF INVENTION: EXPRESSION USING SAID PROMOTER  
; NUMBER OF SEQUENCES: 75  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.  
STREET: 419 7th Street N.W., Ste. 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/913,842  
FILING DATE:

CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 07-073043  
FILING DATE: 30-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/00777  
FILING DATE: 26-MAR-1996

ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: OHBA-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-5197  
TELEFAX: (202) 737-3528  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1406 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-913-842-6

Query Match 1.4%; Score 17; DB 3; Length 1406;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 71 aaaaagaaatacacaca 87  
|||||  
Db 1355 AAAAAGAAATACACA 1371

RESULT 3

US-08-206-176-1  
; Sequence 1, Application US/08206176  
; Patent No. 5639940  
; GENERAL INFORMATION:

; APPLICANT: Garner, Ian  
; APPLICANT: Dairymple, Michael A  
; APPLICANT: Prunkard, Donna E  
; APPLICANT: Foster, Donald C  
; TITLE OF INVENTION: Production of Fibrinogen in Transgenic  
; TITLE OF INVENTION: Animals  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: ZymoGenetics, Inc.  
STREET: 4225 Roosevelt Way, N.E.  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98105

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/206,176  
FILING DATE:

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

NAME: Parker, Gary E  
REGISTRATION NUMBER: 31-648  
REFERENCE/DOCKET NUMBER: 93-15  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-547-8080 ext 322  
TELEFAX: 206-548-2329  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 5943 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
IMMEDIATE SOURCE:  
CLONE: Human Fibrinogen A-alpha chain  
FEATURE:

NAME/KEY: CDS  
LOCATION: join(31..84, 1154..1279, 1739..1922, 3055..3200,  
LOCATION: 3786..5210)  
US-08-206-176-1

Query Match 1.4%; Score 17; DB 1; Length 5943;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 atccctgaaattttaa 123  
|||||  
Db 2055 ATCCCTGAAATTTTAA 2071

RESULT 4

US-08-472-217-1/c  
; Sequence 1, Application US/08472217  
; Patent No. 5726058  
; GENERAL INFORMATION:

; APPLICANT: Alanen-Kurki, Leena  
; APPLICANT: Auvinen, Petri  
; APPLICANT: Jaakkola, Panu  
; APPLICANT: Jalkanen, Markku  
; APPLICANT: Lepp, Sirpa  
; APPLICANT: Mali, Markku  
; APPLICANT: Vihtinen, Tapani  
; APPLICANT: W rri, Anni

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2001, 10:02:38 ; Search time 79.41 Seconds  
(without alignments)  
2896.525 Million cell updates/sec

Title: US-09-389-000-1\_COPY\_735\_1949

Perfect score: 1215

Sequence: 1 ggtctctgaagggtcctt.....ttggcatggtttggtcttg 1215

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 324599 seqs, 94635562 residues

Word size : 0

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

- 1: /cgnl\_7/ptodata/1/ina/5A\_COMB.seq.\*
- 2: /cgnl\_7/ptodata/1/ina/5B\_COMB.seq.\*
- 3: /cgnl\_7/ptodata/1/ina/6A\_COMB.seq.\*
- 4: /cgnl\_7/ptodata/1/ina/6B\_COMB.seq.\*
- 5: /cgnl\_7/ptodata/1/ina/PCTUS\_COMB.seq.\*
- 6: /cgnl\_7/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18	1.5	3442	1 US-08-529-654-3	Sequence 3, Appli
C 2	17	1.4	1406	3 US-08-913-842-6	Sequence 6, Appli
C 3	17	1.4	5943	1 US-08-206-176-1	Sequence 1, Appli
C 4	17	1.4	26700	1 US-08-472-217-1	Sequence 1, Appli
C 5	17	1.4	26700	2 US-08-488-199-5	Sequence 5, Appli
C 6	17	1.4	26700	3 US-08-760-534A-1	Sequence 1, Appli
C 7	16	1.3	422	3 US-08-826-332-15	Sequence 15, Appli
C 8	16	1.3	822	4 US-08-896-164-47	Sequence 47, Appli
C 9	16	1.3	935	1 US-08-162-475A-3	Sequence 3, Appli
C 10	16	1.3	966	1 US-08-162-475A-1	Sequence 1, Appli
C 11	16	1.3	1095	3 US-08-882-501-31	Sequence 31, Appli
C 12	16	1.3	1267	4 US-09-422-869-15	Sequence 15, Appli
C 13	16	1.3	1334	2 US-08-481-658B-44	Sequence 44, Appli
C 14	16	1.3	1334	2 US-08-477-504A-44	Sequence 44, Appli
C 15	16	1.3	1334	2 US-08-486-756A-44	Sequence 44, Appli
C 16	16	1.3	1334	2 US-08-485-862B-44	Sequence 44, Appli
C 17	16	1.3	1334	3 US-08-787-739-44	Sequence 44, Appli
C 18	16	1.3	1334	3 US-08-487-077A-44	Sequence 44, Appli
C 19	16	1.3	1334	3 US-08-485-863A-44	Sequence 44, Appli
C 20	16	1.3	1334	4 US-08-485-049D-44	Sequence 44, Appli
C 21	16	1.3	1554	1 US-08-469-486-1	Sequence 1, Appli
C 22	16	1.3	1554	2 US-08-469-658-1	Sequence 1, Appli
C 23	16	1.3	1635	4 US-09-178-252-10	Sequence 10, Appli
C 24	16	1.3	1770	2 US-08-317-305-1	Sequence 1, Appli
C 25	16	1.3	1770	2 US-08-317-305-3	Sequence 3, Appli
C 26	16	1.3	1770	3 US-08-862-508-1	Sequence 1, Appli
C 27	16	1.3	1770	3 US-08-862-508-3	Sequence 3, Appli

C 28	16	1.3	1770	5 PCT-US95-12508-1	Sequence 1, Appli
C 29	16	1.3	1770	5 PCT-US95-12508-3	Sequence 3, Appli
C 30	16	1.3	1884	3 US-08-784-582-70	Sequence 70, Appli
C 31	16	1.3	2028	2 US-08-933-750C-85	Sequence 85, Appli
C 32	16	1.3	2028	3 US-09-234-613-85	Sequence 85, Appli
C 33	16	1.3	2086	3 US-08-589-028-9	Sequence 9, Appli
C 34	16	1.3	2086	3 US-08-784-582-9	Sequence 9, Appli
C 35	16	1.3	2086	4 US-08-785-271-9	Sequence 9, Appli
C 36	16	1.3	2181	2 US-08-737-371A-1	Sequence 1, Appli
C 37	16	1.3	2181	5 PCT-US95-05853-1	Sequence 1, Appli
C 38	16	1.3	2356	3 US-08-784-582-72	Sequence 72, Appli
C 39	16	1.3	2771	1 US-08-273-411-5	Sequence 5, Appli
C 40	16	1.3	2837	4 US-09-156-316-11	Sequence 11, Appli
C 41	16	1.3	2991	3 US-08-872-094-1	Sequence 1, Appli
C 42	16	1.3	3401	1 US-08-218-943-4	Sequence 4, Appli
C 43	16	1.3	3885	3 US-08-872-094-9	Sequence 9, Appli
C 44	16	1.3	4049	1 US-08-162-809-17	Sequence 17, Appli
C 45	16	1.3	4097	1 US-08-162-809-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-08-529-654-3/c

; Sequence 3, Application US/08529654

; Patent No. 5739284

; GENERAL INFORMATION:

; APPLICANT: HEDIGER, MATTHIAS

; APPLICANT: KANAI, YOSHIKATSU

; TITLE OF INVENTION: COMPOSITIONS CORRESPONDING TO A HIGH

; TITLE OF INVENTION: AFFINITY GLUTAMATE TRANSPORTER MOLECULE AND METHODS FOR

; TITLE OF INVENTION: MAKING AND USING SAME

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.

; STREET: 600 ATLANTIC AVENUE

; CITY: BOSTON

; STATE: MASSACHUSETTS

; COUNTRY: USA

; ZIP: 02210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/529,654

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/194,719

; FILING DATE: 10-FEB-1994

; APPLICATION NUMBER: US 07/965,676

; FILING DATE: 19-OCT-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: JANIUK, ANTHONY J.

; REGISTRATION NUMBER: 29,809

; REFERENCE/DOCKET NUMBER: B0801/7021

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-720-3500

; TELEFAX: 617-720-2441

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3442 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: LEPORIDAE (RABBIT)





PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.5%; Score 18; DB 21; Length 1857;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 403 gatctctctcttcattc 420  
 |||||  
 Db 103 gatctctctcttcattc 120

RESULT 14  
 AAA27120  
 ID AAA27120 standard; cdna; 1868 BP.  
 XX

AC AAA27120;  
 XX  
 DT 08-AUG-2000 (first entry)  
 XX  
 DE cDNA for wheat NADP-specific glutamate dehydrogenase.  
 XX  
 KW NADP-specific glutamate dehydrogenase; herbicide;  
 KW ammonium assimilation protein; altered expression level; wheat;  
 KW nicotinamide adenine dinucleotide phosphate; ss.  
 XX  
 OS Triticum aestivum.  
 XX  
 FH Key Location/Qualifiers  
 CDS 2..1489  
 FT /\*tag= a  
 FT /partial  
 FT /product= "NADP-specific glutamate dehydrogenase"  
 XX  
 PN WO200028006-A2.  
 XX  
 PD 18-MAY-2000.  
 XX  
 PF 04-NOV-1999; 99WO-US25949.  
 XX  
 PR 05-NOV-1998; 98US-0107274.  
 XX  
 PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX  
 PI Cahoon EB, Cahoon RE, Falco SC, Famodu OO, Orozco EM;  
 XX  
 DR WPI: 2000-376524/32.  
 DR P-PSDB: AAY94440.  
 XX  
 PT Novel genes which encode plant NADP-specific glutamate dehydrogenase  
 PT polypeptides useful for producing transgenic plants and useful as  
 PT probes or primers  
 XX  
 PS Claim 2; Page 37-38; 40pp; English.  
 XX  
 CC cDNA libraries from corn, wheat, rice and soybean were used to  
 CC identify cDNA clones encoding ammonium assimilation proteins. The  
 CC clones found represent the first corn, wheat, rice and soybean  
 CC sequences encoding nicotinamide adenine dinucleotide phosphate  
 CC (NADP)-specific glutamate dehydrogenase. The DNA sequences can be  
 CC used as a probe to isolate sequences encoding homologous proteins or as  
 CC a PCR primer to amplify homologous sequences. The nucleic acid fragments  
 CC may be used to create transgenic plants in which the polypeptides are  
 CC present at higher or lower levels than normal. The polypeptides may be  
 CC used to prepare antibodies to these proteins. Additionally, as the  
 CC DNA sequences encode proteins that are involved in a critical step in  
 CC ammonium assimilation, they may be useful in the design of herbicides.  
 CC The nucleic acid fragments may be used as probes for genetically and  
 CC physically mapping the genes for these proteins. The present sequence  
 CC is cDNA for wheat NADP-specific glutamate dehydrogenase.  
 XX  
 SQ Sequence 1868 BP; 523 A; 343 C; 470 G; 532 T; 0 other;

Query Match 1.5%; Score 18; DB 21; Length 1868;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 537 ttgattatgtgaaatata 554  
 |||||  
 Db 990 ttgattatgtgaaatata 1007

RESULT 15  
 AAT72955  
 ID AAT72955 standard; DNA; 2309 BP.  
 XX  
 AC AAT72955;  
 XX

Db 12 gatctctcttcttcattc 29  
RESULT 13  
AAC53975  
ID AAC53975 standard; DNA; 1857 BP.  
AC AAC53975;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 76251.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.



PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160788.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 1.5%; Score 18; DB 21; Length 1766;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 403 gatctctctcttcattc 420  
|||||



```

XX DE PC-gene 2 putative eukaryotic promoter element.
XX XX
XX KW Glyceraldehyde-3-phosphate dehydrogenase; GAPDH; carotenoid;
XX KW synthesis; promoter; recombinant DNA; astaxanthin; ribosomal protein;
XX KW food colouring; ds.
XX OS
XX PH Phaffia rhodozyma.
XX XX
XX XX Key Location/Qualifiers
XX FT promoter 1..385
XX FT /*tag= a
XX FT TATA_signal 249..263
XX FT /*tag= b
XX FT /*note= "putative"
XX FT misc_signal 287..302
XX FT /*tag= c
XX FT /*function= cap-signal
XX FT /*note= "putative"
XX FT misc_RNA 386..388
XX FT /*tag= d
XX FT /*function= start_of_CDS
XX FT misc_feature 85
XX FT /*tag= e
XX FT /*note= "uncertain"
XX FT
XX PN W09723633-A1.
XX XX
XX XX 03-JUL-1997.
XX XX
XX PF 23-DEC-1996; 96WO-EP05887.
XX XX
XX PR 11-APR-1996; 96EP-0200943.
XX PR 22-DEC-1995; 95EP-0203620.
XX XX
XX PA (KONN ) GIST-BROCADES BV.
XX PA (OOIJ/) OOIJEN A J J.
XX XX
XX PI Verdoes JC, Wery J;
XX XX WPI; 1997-351068/32.
XX DR
XX XX Phaffia derived GAPDH and carotenoid synthesis genes and promoter
XX FT fragment - used in the recombinant production of therapeutically
XX FT useful proteins e.g. carotenoids for use in food colouring
XX XX
XX PS Example 4; Page 47; 118pp; English.
XX XX
XX CC The present sequence represents a PC-gene 2 putative eukaryotic
XX CC promoter element, which is mentioned in example 4 of the present
XX CC invention describing novel recombinant DNA. The recombinant DNA
XX CC comprises a transcription promoter operably linked to a downstream
XX CC sequence to be expressed, where the transcription promoter comprises a
XX CC region found upstream of the open reading frame (ORF) of a highly
XX CC expressed Phaffia gene (preferably GAPDH, a ribosomal protein gene, an
XX CC enzyme involved in the biosynthesis pathway). The recombinant DNA can
XX CC be used to transform hosts, preferably Phaffia. These transformed hosts
XX CC are then used in the recombinant production of GAPDH or an enzyme
XX CC involved in carotenoid synthesis, preferably astaxanthin. They may also
XX CC be used to produce a pharmaceutical product. Purified carotenoids can
XX CC be used as colourants in food and/or feed, and also in cosmetics.
XX XX
XX SQ Sequence 388 BP; 97 A; 121 C; 77 G; 92 T; 1 other;

Query Match 1.5%; Score 18; DB 18; Length 388;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 400 aaagatctctctcttcca 417
    |||||
Db 312 aaagatctctcttcca 329

```

RESULT 10  
AAC93495/c  
ID AAC93495 standard; CDNA; 1408 BP.  
XX AAC93495;  
XX  
XX 16-FEB-2001 (first entry)  
XX  
XX Human secreted protein gene 17 SEQ ID NO:27.  
DE  
XX Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KW neotropic; neuroprotective; antibacterial; virucide; fungicide;  
KW ophthalmological; vulnery; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorders; cancer; cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; nervous system disorder;  
KW Alzheimer's disease; ocular disorder; wound healing; skin aging; ss.  
XX  
XX Homo sapiens.  
OS  
XX WO200061626-A1.  
PN  
XX  
XX 19-OCT-2000.  
PD  
XX 06-APR-2000; 2000WO-US09066.  
XX PF  
XX 09-APR-1999; 99US-0128698.  
XX PR 20-JAN-2000; 2000US-0176926.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX PA (ROSE/) ROSEN C A.  
XX  
XX Rosen CA, Ruben SM, Komatsoulis G;  
PI  
XX WPI; 2000-619227/59.  
DR P-PSDB; AAB51843.  
XX  
XX New nucleic acid molecules encoding 49 human secreted proteins for  
PT diagnosing, preventing or ameliorating medical conditions and used for  
PT food additives or preservatives -  
XX  
XX Claim 1; Page 443; 516pp; English.  
PS  
XX Polynucleotide sequences AAC93479 - AAC93527 represent cDNA encoding  
CC human secreted proteins AAB51827 - AAB51875. Sequences AAB51876 -  
CC AAB51927 represent alternative polypeptides encoded by the genes, and  
CC amino acid sequences with which they share homology. The genes and  
CC proteins have activities dependent on the tissues and cells in which they  
CC are expressed. Examples of their activities include immunosuppressive;  
CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;  
CC vasotropic; cerebroprotective; neotropic; neuroprotective; antibacterial;  
CC virucide; fungicide; ophthalmological; and vulnery. The secreted  
CC proteins, polynucleotides, antagonists and agonists may be useful in  
CC treating, preventing and/or diagnosing diseases and disorders such as  
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
CC infections caused by bacteria, viruses and fungi and ocular disorders  
CC e.g. corneal infection. the polypeptides can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin aging due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities, fat content, lipid, protein,  
CC carbohydrate, vitamins, minerals, cofactors and other nutritional  
CC components. Oligonucleotides AAC93470 - AAC93478 and peptide AAB51826 are  
CC used in the isolation and characterisation of the proteins and  
CC polynucleotides of the invention.  
XX  
XX Sequence 1408 BP; 399 A; 254 C; 296 G; 459 T; 0 other;

DT 21-AUG-2000 (first entry)  
 DE Human secreted expressed sequence tag SEQ ID NO:2176.  
 XX  
 XX Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;  
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
 KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;  
 KW antiviral; antidiabetic; antiasthmatic; vulnery; antiparkinsonian;  
 KW anticancer; osteopathic; neuroprotective; nootropic; antipsoriatic;  
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
 KW vaccine; autoimmune diabetes; multiple sclerosis; allergic condition;  
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
 KW central nervous system disorder; Alzheimer's disease; stroke;  
 KW Parkinson's disease; Huntington's disease; coagulation disorder;  
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
 KW tumour; infection; depression; psoriasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200021991-A1.  
 PN  
 XX  
 XX 20-APR-2000.  
 PD  
 XX  
 XX 15-OCT-1999; 99WO-US24206.  
 PF  
 XX  
 XX 15-OCT-1998; 98US-0104436.  
 PR  
 XX  
 XX (GEMY ) GENETICS INST INC.  
 PA  
 XX  
 XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Bowman MR;  
 PI  
 DR WPI: 2000-317938/27.  
 DR  
 XX Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (seSTs), useful for treating various disorders  
 PT such as autoimmune, infectious, and central nervous system disorders -  
 PT  
 XX Claim 1; Page 742; 803pp; English.  
 PS  
 XX AAA43426 to AAA45925 represent specifically claimed secreted expressed  
 CC sequence tags (seSTs), isolated from human, mouse, chicken and rat  
 CC tissue sources. The seSTs can have a range of activities depending on  
 CC the tissues they were isolated from. The activities include:  
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
 CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;  
 CC antiasthmatic; vulnery; anticancer; osteopathic; neuroprotective;  
 CC nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
 CC anticonvulsant; and antidepressant. The seSTs can be used for gene  
 CC therapy and in vaccines. The seSTs are useful as probes for the  
 CC identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the seSTs. Proteins encoded by the seSTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
 CC in the exemplification of the present invention.  
 XX  
 SQ Sequence 122 BP; 35 A; 13 C; 26 G; 46 T; 2 other;

Query Match 1.6%; Score 19; DB 21; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 9.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 aaagaaaaatacacattca 91  
 Db 80 AAAGAAAAATACACATTCA 62  
 RESULT 8  
 AAX58987/c  
 ID AAX58987 standard; cDNA; 6218 BP.  
 XX AAX58987;  
 AC  
 XX 23-AUG-1999 (first entry)  
 DT  
 XX Human transcription regulator MOP8 cDNA.  
 DE  
 XX MOP8; member of the PAS superfamily; bHLH-PAS; human;  
 KW transcription regulator; circadian rhythm; signal transduction; ss.  
 KW  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH 123..3863  
 FT CDS /\*tag= a  
 FT  
 XX WO9928464-A2.  
 PN  
 XX 10-JUN-1999.  
 PD  
 XX 27-NOV-1998; 98WO-US25314.  
 PF  
 XX 28-NOV-1997; 97US-0066863.  
 PR  
 XX (WISC ) WISCONSIN ALUMNI RES FOUND.  
 PA  
 XX Bradfield CA, Gu YZ, Hogenesch JB;  
 PI  
 XX WPI: 1999-371120/31.  
 DR P-PSDB: AAY06296.  
 DR  
 XX Developmental signal transduction associated proteins  
 PT  
 XX Claim 6; Page 101-104; 106pp; English.  
 PS  
 XX This is the nucleotide sequence of a cDNA encoding MOP8 (see  
 CC AAY06296), a novel member of the PAS superfamily, where PAS stands  
 CC for PER/ARNT/SIM domains. MOP8 is thought to be involved in  
 CC circadian rhythm. It shows sequence similarity to other genes  
 CC involved in the circadian pathway, including human PER, Drosophila  
 CC PER and human RIGUI. The invention provides novel MOP 2-9 nucleic  
 CC acids (see AAX58981-88) and proteins (see AAY06289-97). These are  
 CC useful in a variety of research, diagnostic and therapeutic  
 CC applications. Several of the MOPs are alpha-class hypoxia-inducible  
 CC factors. Others are involved in circadian signal transduction.  
 XX  
 SQ Sequence 6218 BP; 1537 A; 1667 C; 1557 G; 1457 T; 0 other;  
 Query Match 1.6%; Score 19; DB 20; Length 6218;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 609 gcaacatgaggtttttaa 627  
 Db 5064 GCAACATGAGGTTTGTAA 5046  
 RESULT 9  
 AAT72956  
 ID AAT72956 standard; DNA; 388 BP.  
 XX  
 AC AAT72956;  
 XX  
 DT 10-MAR-1998 (first entry)

CC prostate and other types of cancer. The invention provides  
 CC diagnostic and therapeutic methods useful in the management of  
 CC various cancers which express PHELiX, including prostate cancer,  
 CC bladder cancer, ovarian cancer and testicular cancer.

SQ Sequence 23 BP; 2 A; 7 C; 5 G; 9 T; 0 other;

Query Match 1.9%; Score 23; DB 21; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.084;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 470 ctgcgtactcttgcgtatgt 492  
 |||||  
 Db 1 ctgcgtactcttgcgtatgt 23

RESULT 5  
 AA294283  
 ID AA294283 standard; DNA; 23 BP.  
 XX  
 AC AA294283;  
 XX  
 DT 03-JUL-2000 (first entry)  
 XX  
 DE PHELiX gene PCR primer 22P4G9.1.  
 XX

PHELiX; human; testis-specific; transcription factor;  
 KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
 KW gene therapy; diagnosis; vaccine; PCR primer;  
 KW chromosome 13q13.1-13.3; ss.

XX Homo sapiens.

OS  
 XX WO200012709-A2.  
 PN  
 XX  
 XX  
 PD 09-MAR-2000.

PF 31-AUG-1999; 99WO-US20137.  
 XX

PR 31-AUG-1998; 98US-0098610.  
 PR 31-OCT-1998; 98US-0106524.  
 XX

XX (UROG-) UROGENESYS INC.

PA (AFAR/) AFAR D E.

PA (HUBE/) HUBERT R S.

PA (RAIT/) RAITANO A B.

XX  
 PI Afar DE, Hubert RS, Raitano AB;  
 XX

DR WPI; 2000-237872/20.

XX  
 PT Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -

PS Example 9; Page 38; 62pp; English.

XX  
 CC The present sequence is that of PCR primer 22P4G9.1, which was used  
 CC for chromosomal mapping of the PHELiX gene (see AA294275) to human  
 CC chromosome 13q.13.1-13.3. PHELiX (see AAY79269) is a transcription  
 CC factor that is normally expressed only in testis tissue, but which  
 CC is up-regulated in prostate and other types of cancer. The  
 CC invention provides diagnostic and therapeutic methods useful in the  
 CC management of various cancers which express PHELiX, including  
 CC prostate cancer, bladder cancer, ovarian cancer and testicular  
 CC cancer.

SQ Sequence 23 BP; 2 A; 7 C; 5 G; 9 T; 0 other;

Query Match 1.9%; Score 23; DB 21; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.084;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 470 ctgcgtactcttgcgtatgt 492  
 |||||  
 Db 1 ctgcgtactcttgcgtatgt 23

RESULT 6  
 AAA54515/c  
 ID AAA54515 standard; DNA; 609 BP.  
 XX  
 AC AAA54515;  
 XX  
 DT 11-APR-2001 (first entry)  
 XX  
 DE Fructan exohydrolase (FEH) partial sequence.

XX Fructan exohydrolase; FEH; transgenic plant; recombination;  
 KW transgene; gene expression; detergent; detergent additive;  
 KW oral care composition; ds.

XX Triticum aestivum.

OS  
 XX WO2000068402-A1.  
 PN  
 XX  
 XX 16-NOV-2000.

XX 08-MAY-2000; 2000WO-EP04226.

XX 06-MAY-1999; 99BE-0000329.

XX (LEUV-) LEUVEN RES & DEV.

XX Van Den Ende W, Van Laere A, De Roover J, Michiels A;  
 PI  
 XX WPI; 2001-007401/01.

XX Novel DNA molecules encoding enzymes having fructan exohydrolase

PT activity for use in transgenic plant production, dental care  
 PT compositions, and in detergents  
 XX  
 XX Claim 11; Fig 6-1; 45pp; English.

XX Transgenic plants such as Cichorium intybus, Cynara scolymus,  
 CC Helianthus tuberosus, Scorzonera hispanica, Oryza sativa, Zea mays,  
 CC Triticum aestivum, Triticum durum, Hordeum vulgare, Secale cereale,  
 CC Avena sativa, Sorghum vulgare, Phleum pratense, Lolium temulentum,  
 CC Dactylis glomerata, Pennisetum americanum, Allium cepa, Agave  
 CC americanum, Agave azul tequilana, Sorghum bicolor and Panicum millaceum,  
 CC transformed with a vector encoding a fructan exohydrolase (FEH)  
 CC enzyme are useful for the recombinant production of FEH or other  
 CC polypeptides having FEH activity. The FEH polypeptides produced  
 CC are useful in detergents or as a detergent additive and in oral  
 CC care compositions.

XX Sequence 609 BP; 159 A; 121 C; 164 G; 165 T; 0 other;

Query Match 1.6%; Score 20; DB 22; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 3.3;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1148 caagaatgatttgaagc 1167  
 |||||  
 Db 119 CAAGAATGATTTCGAAGC 100

RESULT 7  
 AAA45601/c  
 ID AAA45601 standard; cDNA; 122 BP.  
 XX  
 AC AAA45601;

XX

PA (RAIT/) RAITANO A B.  
 XX Afar DE, Hubert RS, Raitano AB;  
 PI WPI: 2000-237872/20.  
 DR  
 XX Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -  
 XX  
 PS Example 1; Page 34; 62pp; English.  
 XX  
 CC The present sequence is that of a primer used in RT-PCR analysis  
 CC to determine expression levels of the human PHELIx gene (see also  
 CC AA294275). Examination of 1st strand cDNAs from 16 normal human  
 CC tissues detected expression of the PHELIx gene only in testis  
 CC tissue. PHELIx expression was also detected in LAPC-4 xenografts.  
 CC PHELIx (see AAY79269) is a transcription factor that is normally  
 CC expressed only in testis tissue, but which is up-regulated in  
 CC prostate and other types of cancer. The invention provides  
 CC diagnostic and therapeutic methods useful in the management of  
 CC various cancers which express PHELIx, including prostate cancer,  
 CC bladder cancer, ovarian cancer and testicular cancer.  
 XX  
 SQ Sequence 24 BP; 2 A; 3 C; 7 G; 12 T; 0 other;

Query Match 2.0%; Score 24; DB 21; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 627 agaaacacacacacccattgagc 650  
 |||||  
 Db 24 AGAAACACACACACCCATTGAGC 1

RESULT 3  
 AA294284/C  
 ID AA294284 standard; DNA; 24 BP.  
 XX  
 AC AA294284;  
 XX  
 DT 03-JUL-2000 (first entry)  
 XX  
 DE PHELIx gene PCR primer 22P4G9.2.  
 XX  
 KW PHELIx; human; testis-specific; transcription factor;  
 KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
 KW gene therapy; diagnosis; vaccine; PCR primer;  
 KW chromosome 13q13.1-13.3; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200012709-A2.  
 XX  
 PD 09-MAR-2000.  
 XX  
 PF 31-AUG-1999; 99WO-US20137.  
 XX  
 PR 31-AUG-1998; 98US-0098610.  
 PR 31-OCT-1998; 98US-0106524.  
 XX  
 PA (UROC-) UROGENESYS INC.  
 PA (AFAR/) AFAR D E.  
 PA (HUBE/) HUBERT R S.  
 PA (RAIT/) RAITANO A B.  
 XX  
 PI Afar DE, Hubert RS, Raitano AB;  
 XX  
 DR WPI: 2000-237872/20.  
 XX  
 PT Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder

PT and ovarian tumors -  
 XX  
 PS Example 9; Page 38; 62pp; English.  
 XX  
 CC The present sequence is that of PCR primer 22P4G9.2, which was used  
 CC for chromosomal mapping of the PHELIx gene (see AA294275) to human  
 CC chromosome 13q.13.1-13.3. PHELIx (see AAY79269) is a transcription  
 CC factor that is normally expressed only in testis tissue, but which  
 CC is up-regulated in prostate and other types of cancer. The  
 CC invention provides diagnostic and therapeutic methods useful in the  
 CC management of various cancers which express PHELIx, including  
 CC prostate cancer, bladder cancer, ovarian cancer and testicular  
 CC yancer.  
 XX  
 SQ Sequence 24 BP; 2 A; 3 C; 7 G; 12 T; 0 other;

Query Match 2.0%; Score 24; DB 21; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 627 agaaacacacacccattgagc 650  
 |||||  
 Db 24 AGAAACACACACACCCATTGAGC 1

RESULT 4  
 AA294281  
 ID AA294281 standard; DNA; 23 BP.  
 XX  
 AC AA294281;  
 XX  
 DT 03-JUL-2000 (first entry)  
 XX  
 DE PHELIx gene PCR primer.  
 XX  
 KW PHELIx; human; testis-specific; transcription factor;  
 KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
 KW gene therapy; diagnosis; vaccine; PCR primer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200012709-A2.  
 XX  
 PD 09-MAR-2000.  
 XX  
 PF 31-AUG-1999; 99WO-US20137.  
 XX  
 PR 31-AUG-1998; 98US-0098610.  
 PR 31-OCT-1998; 98US-0106524.  
 XX  
 PA (UROC-) UROGENESYS INC.  
 PA (AFAR/) AFAR D E.  
 PA (HUBE/) HUBERT R S.  
 PA (RAIT/) RAITANO A B.  
 XX  
 PI Afar DE, Hubert RS, Raitano AB;  
 XX  
 DR WPI: 2000-237872/20.  
 XX  
 PT Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -  
 XX  
 PS Example 1; Page 34; 62pp; English.  
 XX  
 CC The present sequence is that of a primer used in RT-PCR analysis  
 CC to determine expression levels of the human PHELIx gene (see also  
 CC AA294275). Examination of 1st strand cDNAs from 16 normal human  
 CC tissues detected expression of the PHELIx gene only in testis  
 CC tissue. PHELIx expression was also detected in LAPC-4 xenografts.  
 CC PHELIx (see AAY79269) is a transcription factor that is normally  
 CC expressed only in testis tissue, but which is up-regulated in  
 CC prostate and other types of cancer. The invention provides  
 CC diagnostic and therapeutic methods useful in the management of  
 CC various cancers which express PHELIx, including prostate cancer,  
 CC bladder cancer, ovarian cancer and testicular cancer.

PI Afar DE, Hubert RS, Raitano AB;  
 XX WPI: 2000-237872/20.  
 DR P-PSDB; AAY79269.  
 XX  
 XX Testis specific Helix Loop Helix proteins expressed in cancers and  
 PT useful for the prevention, diagnosis and treatment of prostate, bladder  
 PT and ovarian tumors -  
 XX  
 XX Claim 5; Fig 2A-D; 62pp; English.  
 XX  
 XX This sequence is that of cDNA encoding human PHELIX (see AAY79269),  
 CC a novel basic Helix Loop Helix protein thought to act as a  
 CC transcription factor. PHELIX normally exhibits a testis-specific  
 CC expression pattern but is up-regulated in prostate and other types  
 CC of cancer. The cDNA clone, termed GTPIC12 (ATCC 98956), was  
 CC isolated from a normal testis cDNA library using a suppression  
 CC subtractive hybridization method. The gene maps to chromosome  
 CC 13q13.1-13.3. The invention provides diagnostic and therapeutic  
 CC methods useful in the management of various cancers which express  
 CC PHELIX, including prostate cancer, bladder cancer, ovarian cancer  
 CC and testicular cancer, including therapies aimed at inhibition the  
 CC transcription, translation, processing or function of PHELIX. The  
 CC expression pattern of PHELIX suggests that is an ideal target for a  
 CC cancer vaccine approach to prostate cancer.  
 XX  
 XX Sequence 2128 BP; 675 A; 463 C; 464 G; 526 T; 0 other;  
 SQ

Query Match 100.0%; Score 1215; DB 21; Length 2128;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggttctctgaagtgctcttcaactaagtccgagagctggaagccatcaagttta 60  
 Db  
 735 ggttctctgaagtgctcttcaactaagtccgagagctggaagccatcaagttta 794  
 Qy 61 tagatttgcaaaagaaaaatacacattcaactgttgtttttataatccctgaaaaattt 120  
 Db 795 tagatttgcaaaagaaaaatacacattcaactgttgtttttataatccctgaaaaattt 854  
 Qy 121 taaagttgtatttcaggagcgaatgaatgattgcttcaactgaaccactgacaatgga 180  
 Db 855 taaagttgtatttcaggagcgaatgaatgattgcttcaactgaaccactgacaatgga 914  
 Qy 181 aaaaatgagtaattgtgttaataactggacacacatgtccctcaaacactgttaagactga 240  
 Db 915 aaaaatgagtaattgtgttaataactggacacacatgtccctcaaacactgttaagactga 974  
 Qy 241 aaacgcaactgggctgaagaacttgattgcccctgcagaggttccctacagcgacaacct 300  
 Db 975 aaacgcaactgggctgaagaacttgattgcccctgcagaggttccctacagcgacaacct 1034  
 Qy 301 gggatatttccctactgactattgctcgtcctgaatctttaaggaaatggcaatggct 360  
 Db 1035 gggatatttccctactgactattgctcgtcctgaatctttaaggaaatggcaatggct 1094  
 Qy 361 tgaattaaatgcttcgttgcagagttcagaaaaaacaagaatctctcttcttcattc 420  
 Db 1095 tgaattaaatgcttcgttgcagagttcagaaaaaacaagaatctctcttcttcattc 1154  
 Qy 421 aagcaagaaaaactaagaaaggaagaatgatcgccgttcagttcttctgaggaacagttga 480  
 Db 1155 aagcaagaaaaactaagaaaggaagaatgatcgccgttcagttcttctgaggaacagttga 1274  
 Qy 481 cttgccgtatgtaaaaggaagaagaatgatcgccgttcagttcttctgaggaacagttga 540  
 Db 1215 cttgccgtatgtaaaaggaagaagaatgatcgccgttcagttcttctgaggaacagttga 1274  
 Qy 541 ttatgtgaatatatccgggagaaaaatctctccagcgcgttatgcccagattacagaagc 600  
 Db 1275 ttatgtgaatatatccgggagaaaaatctctccagcgcgttatgcccagattacagaagc 1334

Qy 601 acttcagagcaacatgaggttttgaagaaacaaacacccattgagctgtctctccc 660  
 Db 1335 acttcagagcaacatgaggttttgaagaaacaaacacccattgagctgtctctccc 1394  
 Qy 661 aggcactgtcatggcacagcgggaaacacagtgatgagcacttactccctcgagagagg 720  
 Db 1395 aggcactgtcatggcacagcgggaaacacagtgatgagcacttactccctcgagagagg 1454  
 Qy 721 gctccaattctctgactaatacgtgtctggaatgggtgctccactcctgatgcagagagctc 780  
 Db 1455 gctccaattctctgactaatacgtgtctggaatgggtgctccactcctgatgcagagagctc 1514  
 Qy 781 cttggatgaagctgtgagagttccatcaagctccgcctcagagaatgctattgtgtatcc 840  
 Db 1515 cttggatgaagctgtgagagttccatcaagctccgcctcagagaatgctattgtgtatcc 1574  
 Qy 841 atataaaactcacatttccagtcagcgtctctctgaattccttgatactgacagata 900  
 Db 1575 atataaaactcacatttccagtcagcgtctctctgaattccttgatactgacagata 1634  
 Qy 901 ttattctaaagtcaccccttctctacgatgcaactgctgtaacaaatcagaaacatttcaat 960  
 Db 1635 ttattctaaagtcaccccttctctacgatgcaactgctgtaacaaatcagaaacatttcaat 1694  
 Qy 961 tcatttaacttcagccatgccccgggtctcaagcttctccctcgccactgcacttctggg 1020  
 Db 1695 tcatttaacttcagccatgccccgggtctcaagcttctccctcgccactgcacttctggg 1754  
 Qy 1021 ttggccagacgtgcactacacatcccaactgctgcaacagttttggcgctattaaaag 1080  
 Db 1755 ttggccagacgtgcactacacatcccaactgctgcaacagttttggcgctattaaaag 1814  
 Qy 1081 cacatgtttgaaatttcacactctcaacacactactggtggtgagtttgacaatctaggaaa 1140  
 Db 1815 cacatgtttgaaatttcacactctcaacacactactggtggtgagtttgacaatctaggaaa 1874  
 Qy 1141 agtggaaacaaagaatgattttgaaagctcccaacaaagacctaataataaaagattggc 1200  
 Db 1875 agtggaaacaaagaatgattttgaaagctcccaacaaagacctaataataaaagattggc 1934  
 Qy 1201 atggtttggcttctg 1215  
 Db 1935 atggtttggcttctg 1949

RESULT 2  
 AAZ94282/C  
 ID AAZ94282 standard; DNA; 24 BP.  
 XX  
 AC AAZ94282;  
 XX  
 DT 03-JUL-2000 (first entry)  
 XX  
 DE PHELIX gene PCR primer.  
 XX  
 KW PHELIX; human; testis-specific; transcription factor;  
 KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
 KW gene therapy; diagnosis; vaccine; PCR primer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200012709-A2.  
 XX  
 XX 09-MAR-2000.  
 XX  
 XX 31-AUG-1999; 99WO-US20137.  
 XX  
 PR 31-AUG-1998; 98US-0098610.  
 PR 31-OCT-1998; 98US-0106524.  
 XX  
 XX (UROC-) UROGENESYS INC.  
 PA (AFAR/) AFAR D E.  
 PA (HUBE/) HUBERT R S.

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2001, 10:31:28 : Search time 125.9 Seconds  
(without alignments)  
6059.575 Million cell updates/sec

Title: us-09-389-000-1\_copy\_735\_1949  
Perfect score: 1215  
Sequence: 1 ggtctctgaaggcctt.....ttggcatggtttggtctctg 1215

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 730101 seqs, 313950809 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Listing first 45 summaries

- Database : N\_Geneseq\_0601.\*
- 1: /cgnl\_9/gcgdata/geneseq/geneseq/NA1980.DAT.\*
  - 2: /cgnl\_9/gcgdata/geneseq/geneseq/NA1981.DAT.\*
  - 3: /cgnl\_9/gcgdata/geneseq/geneseq/NA1982.DAT.\*
  - 4: /cgnl\_9/gcgdata/geneseq/geneseq/NA1983.DAT.\*
  - 5: /cgnl\_9/gcgdata/geneseq/geneseq/NA1984.DAT.\*
  - 6: /cgnl\_9/gcgdata/geneseq/geneseq/NA1985.DAT.\*
  - 7: /cgnl\_9/gcgdata/geneseq/geneseq/NA1986.DAT.\*
  - 8: /cgnl\_9/gcgdata/geneseq/geneseq/NA1987.DAT.\*
  - 9: /cgnl\_9/gcgdata/geneseq/geneseq/NA1988.DAT.\*
  - 10: /cgnl\_9/gcgdata/geneseq/geneseq/NA1989.DAT.\*
  - 11: /cgnl\_9/gcgdata/geneseq/geneseq/NA1990.DAT.\*
  - 12: /cgnl\_9/gcgdata/geneseq/geneseq/NA1991.DAT.\*
  - 13: /cgnl\_9/gcgdata/geneseq/geneseq/NA1992.DAT.\*
  - 14: /cgnl\_9/gcgdata/geneseq/geneseq/NA1993.DAT.\*
  - 15: /cgnl\_9/gcgdata/geneseq/geneseq/NA1994.DAT.\*
  - 16: /cgnl\_9/gcgdata/geneseq/geneseq/NA1995.DAT.\*
  - 17: /cgnl\_9/gcgdata/geneseq/geneseq/NA1996.DAT.\*
  - 18: /cgnl\_9/gcgdata/geneseq/geneseq/NA1997.DAT.\*
  - 19: /cgnl\_9/gcgdata/geneseq/geneseq/NA1998.DAT.\*
  - 20: /cgnl\_9/gcgdata/geneseq/geneseq/NA1999.DAT.\*
  - 21: /cgnl\_9/gcgdata/geneseq/geneseq/NA2000.DAT.\*
  - 22: /cgnl\_9/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1215	100.0	2128	21	Human PHELI
2	24	2.0	24	21	PHELI gene PCR pr
3	24	2.0	24	21	PHELI gene PCR pr
4	23	1.9	23	21	PHELI gene PCR pr
5	23	1.9	23	21	PHELI gene PCR pr
6	20	1.6	609	22	Fructan exohydrola
7	19	1.6	122	21	Human secreted exp
8	19	1.6	6218	20	Human transcriptio
9	18	1.5	388	18	PC-gene 2 putative
10	18	1.5	1408	21	Human secreted pro
11	18	1.5	1751	21	Human secreted pro

12	18	1.5	1766	21	AA54164	Arabidopsis thalia
13	18	1.5	1857	21	AA53975	Arabidopsis thalia
14	18	1.5	1868	21	AA27120	cDNA for wheat NAD
15	18	1.5	2309	18	AA72955	Ubiquitin-ribosoma
c 16	18	1.5	3442	19	AAV1967	Rabbit excitatory
17	18	1.5	6901	18	AA78853	Human lecithin-cho
18	18	1.5	11427	20	AA13102	Enterococcus faeca
19	17	1.4	28	16	AA086825	Primer #2 for acet
20	17	1.4	170	21	AAFL4771	Aspergillus oryzae
c 21	17	1.4	304	18	AA78693	Staphylococcus aur
c 22	17	1.4	330	21	AAFL1660	Aspergillus niger
23	17	1.4	341	21	AA28532	Human secreted pro
c 24	17	1.4	344	21	AA21590	Human secreted pro
25	17	1.4	370	18	AAV78456	Staphylococcus aur
c 26	17	1.4	400	18	AAV78096	Staphylococcus aur
c 27	17	1.4	403	22	AA93376	cDNA encoding SRT
28	17	1.4	457	21	AA28157	Human purH amplico
c 29	17	1.4	495	21	AA201443	Human secreted pro
30	17	1.4	597	21	AA203992	Human secreted pro
c 31	17	1.4	721	21	AA53953	Arabidopsis thalia
32	17	1.4	1053	18	AA83926	DNA encoding a Sta
33	17	1.4	1339	20	AA242226	Human normal blad
34	17	1.4	1406	17	AA741619	Tomato endo-xylogl
35	17	1.4	1448	21	AA21931	Human breast and o
c 36	17	1.4	1463	20	AA20131	Enterococcus faeca
c 37	17	1.4	1503	21	AA251253	Human RNA-associat
c 38	17	1.4	1685	19	AA59874	DNA encoding a Sta
39	17	1.4	1843	20	AA207724	C. boidinii dlhydr
40	17	1.4	1941	21	AA76885	Human ORFX ORF2440
c 41	17	1.4	2045	21	AAFL16314	Human prostate can
42	17	1.4	2061	20	AA233929	Human PRO329 nucle
43	17	1.4	2061	21	AA278471	Human PRO329 (UNQ2
44	17	1.4	2098	15	AA555320	Recombinant hyalur
45	17	1.4	2099	21	AA27046	Human cell surface

ALIGNMENTS

RESULT	1
AA294275	
ID	AA294275 standard; cDNA; 2128 BP.
XX	AC
AC	AA294275;
DT	03-JUL-2000 (first entry)
DE	Human PHELI cDNA clone GTP1C12.
XX	PHELI; human; testis-specific; transcription factor;
KW	prostate cancer; bladder cancer; ovary cancer; testicular cancer;
KW	gene therapy; diagnosis; vaccine; chromosome 13q13.1-13.3; ss.
XX	Homo sapiens.
OS	
XX	
PH	Key
FT	CDS
FT	733..1953
XX	/*tag= a
PN	WO200012709-A2.
XX	
PD	09-MAR-2000.
XX	
PF	31-AUG-1999; 99WO-US20137.
XX	
PR	31-AUG-1998; 98US-0098610.
XX	
PR	31-OCT-1998; 98US-0106524.
XX	
PA	(UROC-) UROGENESYS INC.
PA	(AFAR) AFAR D E.
PA	(HUBE/) HUBERT R S.
PA	(RAIT/) RAITANO A B.
XX	



**THIS PAGE BLANK (USPTO)**

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 175105)
TITLE	DOE Joint Genome Institute.
JOURNAL	Sequencing of Human Chromosome 16
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 175105)
TITLE	DOE Joint Genome Institute.
JOURNAL	Direct Submission
COMMENT	Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA On Aug 31, 2000 this sequence version replaced gi:8576022. -----Genome Center Center: Joint Genome Institute Center Code: JGI Web site: <a href="http://www.jgi.doe.gov">http://www.jgi.doe.gov</a>

Search completed: September 15, 2001, 11:49:29  
Job time: 6562 sec

```

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 159818: contig of 159818 bp in length.
FEATURES
    source
        Location/Qualifiers
            1..159818
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="16"
            /clone="RP11-223H1"
BASE COUNT 48907 a 30967 c 32591 g 47352 t 1 others
ORIGIN

Query Match 1.7%; Score 21; DB 60; Length 159818;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 299 ctgggatatctctactgat 319
|||||
Db 48198 CTGGGATATTTCTACTGAT 48178

RESULT 13
AC009039/c
LOCUS AC009039 163229 bp DNA PRI 01-AUG-2000
DEFINITION Homo sapiens chromosome 16 clone RP11-152014, complete sequence.
ACCESSION AC009039
VERSION AC009039.6 GI:9625324
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 163229)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
TITLE Unpublished
REFERENCE 2 (bases 1 to 163229)
AUTHORS DOE Joint Genome Institute.
JOURNAL Direct Submission
REFERENCE 3 (bases 1 to 163229)
AUTHORS Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 163229)
AUTHORS Draft Sequence Produced by DOE Joint Genome Institute
JOURNAL Direct Submission
REFERENCE 5 (bases 1 to 163229)
AUTHORS www.jgi.doe.gov
JOURNAL Finishing Completed at Stanford Human Genome Center
COMMENT On Aug 1, 2000 this sequence version replaced gi:6938890.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.5.
Location/Qualifiers
    1..163229
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="16"
    /clone="RP11-152014"
BASE COUNT 49927 a 31684 c 33278 g 48340 t
ORIGIN

Query Match 1.7%; Score 21; DB 87; Length 163229;

```

```

Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 299 ctgggatatctctactgat 319
|||||
Db 51660 CTGGGATATTTCTACTGAT 51640

RESULT 14
AC008526/c
LOCUS AC008526 168243 bp DNA PRI 19-MAY-2000
DEFINITION Homo sapiens chromosome 5 clone CTC-470C15, complete sequence.
ACCESSION AC008526
VERSION AC008526.5 GI:7940321
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 168243)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
TITLE Unpublished
REFERENCE 2 (bases 1 to 168243)
AUTHORS DOE Joint Genome Institute.
JOURNAL Direct Submission
REFERENCE 3 (bases 1 to 168243)
AUTHORS Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 168243)
AUTHORS Draft Sequence Produced by DOE Joint Genome Institute
JOURNAL Direct Submission
REFERENCE 5 (bases 1 to 168243)
AUTHORS www.jgi.doe.gov
JOURNAL Finishing Completed at Stanford Human Genome Center
COMMENT On May 19, 2000 this sequence version replaced gi:7711285.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Quality: Phrap Quality >=40 99.7% of Sequence;
Estimated Total Number of Errors is 0.3.
Location/Qualifiers
    1..168243
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="5"
    /clone="CTC-470C15"
BASE COUNT 54285 a 31638 c 30157 g 52163 t
ORIGIN

Query Match 1.7%; Score 21; DB 86; Length 168243;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 436 aagaagggaagaatcaata 456
|||||
Db 52396 AAGAAGGGAAGAATCAATA 52376

RESULT 15
AC010531/c
LOCUS AC010531 175105 bp DNA HTG 31-AUG-2000
DEFINITION Homo sapiens chromosome 16 clone RP11-17818, WORKING DRAFT
SEQUENCE, 2 ordered pieces.
ACCESSION AC010531
VERSION AC010531.5 GI:9954623
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

```

## JOURNAL

Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 22, 2000 this sequence version replaced gi:6850364.

## COMMENT

All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L661

Center clone name: 92\_A\_1

----- Summary Statistics

Sequencing vector: M13; M77815; 99% of reads

Sequencing vector: Plasmid; n/a; %0.0% of reads

0.507900677200903Chemistry: Dye-terminator Big Dye; 100% of reads

## reads

Assembly program: Phrap; version 0.960731

Consensus quality: 150584 bases at least Q40

Consensus quality: 153858 bases at least Q30

Consensus quality: 155115 bases at least Q20

Insert size: 158000; agarose-fp

Insert size: 156118; sum-of-contigs

Quality coverage: 5.8 in Q20 bases; agarose-fp

Quality cov.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

* 1 1036: contig of 1036 bp in length
* 1037 1136: gap of 100 bp
* 1137 2225: contig of 1089 bp in length
* 2226 2325: gap of 100 bp
* 2326 4045: contig of 1720 bp in length
* 4046 4145: gap of 100 bp
* 4146 9757: contig of 5612 bp in length
* 9758 9857: gap of 100 bp
* 9858 19788: contig of 9931 bp in length
* 19789 19888: gap of 100 bp
* 19889 31050: contig of 11162 bp in length
* 31051 31150: gap of 100 bp
* 31151 40302: contig of 9152 bp in length
* 40303 40402: gap of 100 bp
* 40403 53901: contig of 13499 bp in length
* 53902 54001: gap of 100 bp
* 54002 71033: contig of 17032 bp in length
* 71034 71133: gap of 100 bp
* 71134 89940: contig of 18807 bp in length
* 89941 90040: gap of 100 bp
* 90041 118118: contig of 28078 bp in length
* 118119 118218: gap of 100 bp
* 118219 157218: contig of 39000 bp in length.

```

## FEATURES

## source

```

1. .157218
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="RP11-92A1"
  /clone_lib="RPC1-11 Human Male BAC"

```

## misc\_feature

```

1. .1036
  /note="assembly_fragment"

```

## misc\_feature

```

1137. .2225
  /note="assembly_fragment"

```

## misc\_feature

```

2326. .4045
  /note="assembly_fragment"

```

## misc\_feature

```

4146. .9757
  /note="assembly_fragment"

```

## misc\_feature

```

9858. .19788

```

```

misc_feature
19889. .31050
  /note="assembly_fragment"

```

```

misc_feature
31151. .40302
  /note="assembly_fragment"

```

```

vector_end:SP6
vector_side:left"
40403. .53901
  /note="assembly_fragment"

```

```

54002. .71033
  /note="assembly_fragment"

```

```

71134. .89940
  /note="assembly_fragment"

```

```

clone_end:T7
vector_side:right"
90041. .118118
  /note="assembly_fragment"

```

```

118219. .157218
  /note="assembly_fragment"

```

```

BASE COUNT 49691 a 28915 c 27359 g 50152 t 1101 others
ORIGIN

```

```

Query Match 1.7%; Score 21; DB 65; Length 157218;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 417 attcaagcaggaagaaactaa 437
|||||
Db 8881 ATTCAAGCAGGAAACTAA 8901

```

```

RESULT 12
AC007337/c
LOCUS
DEFINITION
Homo sapiens chromosome 16 clone RP11-223H1, WORKING DRAFT
SEQUENCE
AC007337
AC007337.4 GI:9795552
HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 159818)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 159818)
Bruce, D., Mundt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
Jones, M., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
Bryant, J., Tesmer, J., Melnick, L., Longmire, J., White, S., Tatum, O.,
Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
McMurry, K., Han, C. and Deaven, L.
Direct Submission
Submitted (21-APR-1999) Center for Human Genome Studies, DOE Joint
Genome Institute, Los Alamos National Laboratory, MS M888, Los
Alamos, NM 87545, USA
On Aug 11, 2000 this sequence version replaced gi:9090191.
-----Genome Center
Center: Joint Genome Institute
Center code: JGI
Web site: http://www.jgi.doe.gov
-----

```

```

Sequencing Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

```



```

* 9598 11117: contig of 1520 bp in length
* 11118 11217: gap of 100 bp
* 12118 12816: contig of 1599 bp in length
* 12817 12916: gap of 100 bp
* 12917 14048: contig of 1132 bp in length
* 14049 14149: gap of 100 bp
* 14149 16422: contig of 2274 bp in length
* 16423 16522: gap of 100 bp
* 16523 17466: contig of 944 bp in length
* 17467 17566: gap of 100 bp
* 17567 19094: contig of 1528 bp in length
* 19095 19194: gap of 100 bp
* 19195 20931: contig of 1737 bp in length
* 20932 21031: gap of 100 bp
* 21032 22724: contig of 1693 bp in length
* 22725 22824: gap of 100 bp
* 22825 24816: contig of 1992 bp in length
* 24817 24916: gap of 100 bp
* 24917 27863: contig of 2947 bp in length
* 27864 27963: gap of 100 bp
* 27964 30923: contig of 2966 bp in length
* 30930 31029: gap of 100 bp
* 31030 34191: contig of 3162 bp in length
* 34192 34291: gap of 100 bp
* 34292 36522: contig of 2231 bp in length
* 36523 36622: gap of 100 bp
* 36623 38549: contig of 1927 bp in length
* 38550 38649: gap of 100 bp
* 38650 41768: contig of 3119 bp in length
* 41769 41868: gap of 100 bp
* 41869 44626: contig of 2758 bp in length
* 44627 44726: gap of 100 bp
* 44727 49894: contig of 5168 bp in length
* 49895 49994: gap of 100 bp
* 49995 55860: contig of 5866 bp in length
* 55861 55960: gap of 100 bp
* 55961 61473: contig of 5513 bp in length
* 61474 61573: gap of 100 bp
* 61574 67234: contig of 5661 bp in length
* 67235 67334: gap of 100 bp
* 67335 74758: contig of 7424 bp in length
* 74759 74858: gap of 100 bp
* 74859 83919: contig of 9061 bp in length
* 83920 84019: gap of 100 bp
* 84020 94760: contig of 10741 bp in length
* 94761 94860: gap of 100 bp
* 94861 113489: contig of 18629 bp in length
* 113490 113589: gap of 100 bp
* 113590 133716: contig of 20127 bp in length
* 133717 133816: gap of 100 bp
* 133817 148624: contig of 14808 bp in length.

```

## FEATURES

```

source
  Location/Qualifiers
    1. 148624
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="RP11-788A9"
      /clone_lib="RPC1-11 Human Male BAC"
  misc_feature
    1. 11163
      /note="assembly_fragment"
      1264..1977
      /note="assembly_fragment"
      2078..3067
      /note="assembly_fragment"
      3188..4340
      /note="assembly_fragment"
      4441..5361
      /note="assembly_fragment"
      5462..6605
      /note="assembly_fragment"
      6706..7814
      /note="assembly_fragment"
      7915..9497
      /note="assembly_fragment"

```

```

misc_feature
  9598..11117
  /note="assembly_fragment"
misc_feature
  11218..12816
  /note="assembly_fragment"
misc_feature
  12917..14048
  /note="assembly_fragment"
misc_feature
  14149..16422
  /note="assembly_fragment"
misc_feature
  16523..17466
  /note="assembly_fragment"
misc_feature
  17567..19094
  /note="assembly_fragment"
misc_feature
  19195..20931
  /note="assembly_fragment"
misc_feature
  21032..22724
  /note="assembly_fragment"
misc_feature
  22825..24816
  /note="assembly_fragment"
misc_feature
  24917..27863
  /note="assembly_fragment"
misc_feature
  27964..30929
  /note="assembly_fragment"
misc_feature
  31030..34191
  /note="assembly_fragment"
misc_feature
  34292..36522
  /note="assembly_fragment"
misc_feature
  36623..38549
  /note="assembly_fragment"
misc_feature
  38650..41768
  /note="assembly_fragment"
misc_feature
  41869..44626
  /note="assembly_fragment"
misc_feature
  44727..49894
  /note="assembly_fragment"
misc_feature
  49995..55860
  /note="assembly_fragment"
misc_feature
  55961..61473
  /note="assembly_fragment"
misc_feature
  61574..67234
  /note="assembly_fragment"
misc_feature
  67335..74758

```

Query Match 1.7%; Score 21; DB 77; Length 148624;  
 Best Local Similarity 100.0%; Pred. No. 8.8;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 65.tttggcaaaagaaaataca 85
|||||
Db 85486 TTTGGCAAAAGAAAATACA 85466

```

## RESULT 10

```

AC007670/c
LOCUS AC007670 150070 bp DNA HTG
DEFINITION Homo sapiens clone RP11-14A7, WORKING DRAFT SEQUENCE, 11 unordered
pieces.
ACCESSION AC007670
VERSION AC007670.2 GI:7321636
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 150070)
REFERENCE
  AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE Homo sapiens, clone RP11-14A7
  JOURNAL Unpublished
  REFERENCE
    2 (bases 1 to 150070)
    Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
    Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
    Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
    Cooke,P., DeArellano,K., Depayre,E., Devon,K., Dewar,K.,
    Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,

```

```

* 2398 2497: gap of unknown length
* 3861: contig of 1364 bp in length
* 3961: gap of unknown length
* 5284: contig of 1323 bp in length
* 5384: gap of unknown length
* 5385 6868: contig of 1484 bp in length
* 6869 6968: gap of unknown length
* 6969 9049: contig of 2081 bp in length
* 9050 9149: gap of unknown length
* 9150 10442: contig of 1293 bp in length
* 10443 10542: gap of unknown length
* 10543 13634: contig of 3092 bp in length
* 13635 20114: gap of unknown length
* 20115 20114: contig of 6380 bp in length
* 20115 24952: contig of 4738 bp in length
* 24953 25052: gap of unknown length
* 25053 33127: contig of 8074 bp in length
* 33127 33226: gap of unknown length
* 33227 35908: contig of 2682 bp in length
* 35909 36008: gap of unknown length
* 36009 45728: contig of 9720 bp in length
* 45729 45828: gap of unknown length
* 45829 54267: contig of 8439 bp in length
* 54268 54367: gap of unknown length
* 54368 63147: contig of 8780 bp in length
* 63148 63248: gap of unknown length
* 63248 67622: contig of 4375 bp in length
* 67623 67723: gap of unknown length
* 67723 76298: contig of 8576 bp in length
* 76299 76399: gap of unknown length
* 76399 87565: contig of 11167 bp in length
* 87566 87665: gap of unknown length
* 87666 103770: contig of 16105 bp in length
* 103771 103870: gap of unknown length
* 103871 128024: contig of 24154 bp in length
* 128025 128124: gap of unknown length
* 128125 144793: contig of 16669 bp in length.

FEATURES             Location/Qualifiers
     source            1..144793
     organism="Homo sapiens"
     db_xref="taxon:9606"
     chromosome="5"
     clone="CTB-160022"
     /clone_lib="Caltech human BAC library B"
BASE COUNT          45072 a 27990 c 28289 g 41348 t 2094 others
ORIGIN

Query Match          1.7%; Score 21; DB 60; Length 144793;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 394 aaacaaagaatctctctct 414
|||||
Db 107027 AAACAAAAGAATCTCTCT 107007

RESULT 9
AC084815/c          148624 bp      DNA      HTG      15-JAN-2001
LOCUS              Homo sapiens clone RP11-788A9, WORKING DRAFT SEQUENCE, 34 unordered
DEFINITION          pieces.
AC084815
VERSION             AC084815.2 GI:12229388
KEYWORDS            HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE              human.
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE           1 (bases 1 to 148624)
AUTHORS            Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE              Homo sapiens, clone RP11-788A9

```

Unpublished  
2 (bases 1 to 148624)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L.,  
Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,  
Choepeil,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,  
DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,  
FitZHugh,W., Gage,D., Galagan,J., Gardyna,S., Glnde,S., Goyette,M.,  
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,  
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,  
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,  
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
Willson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zimmer,A. and Zody,M.

Direct Submission  
Submitted (19-NOV-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 15, 2001 this sequence version replaced gi:11225647.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L11509  
Center clone name: 788\_A\_9  
----- Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 134328 bases at least Q40  
Consensus quality: 141261 bases at least Q30  
Consensus quality: 143853 bases at least Q20  
Insert size: 164000; agarose-fp  
Insert size: 145324; sum-of-contigs  
Quality coverage: 3.6 in Q20 bases; agarose-fp  
Quality coverage: 4.0 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 34 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1163: contig of 1163 bp in length  
\* 1164 1263: gap of 100 bp  
\* 1264 1977: contig of 714 bp in length  
\* 1978 2077: gap of 100 bp  
\* 2078 3067: contig of 990 bp in length  
\* 3068 3167: gap of 100 bp  
\* 3168 4340: contig of 1173 bp in length  
\* 4341 4440: gap of 100 bp  
\* 4441 5361: contig of 921 bp in length  
\* 5362 5461: gap of 100 bp  
\* 5462 6605: contig of 1144 bp in length  
\* 6606 6705: gap of 100 bp  
\* 6706 7814: contig of 1109 bp in length  
\* 7815 7914: gap of 100 bp  
\* 7915 9497: contig of 1583 bp in length  
\* 9498 9597: gap of 100 bp

repeat\_region /rpt\_family="(CATA)n"  
753. .900  
repeat\_region /rpt\_family="Alu"  
903. .922  
repeat\_region /rpt\_family="AT-rich"  
923. .1310  
repeat\_region /rpt\_family="L1"  
1311. .1324  
repeat\_region /rpt\_family="AT-rich"  
1811. .1906  
repeat\_region /rpt\_family="MIR"  
2354. .2650  
repeat\_region /rpt\_family="Alu"  
2652. .2707  
repeat\_region /rpt\_family="(TAAA)n"  
2737. .2974  
repeat\_region /rpt\_family="MIR"  
3206. .3513  
repeat\_region /rpt\_family="Alu"  
3524. .3795  
repeat\_region /rpt\_family="Alu"  
3797. .3821  
repeat\_region /rpt\_family="(TAAA)n"  
3953. .4107  
repeat\_region /rpt\_family="(TC)n"  
4119. .4171  
repeat\_region /rpt\_family="NaLR"  
4987. .5070  
repeat\_region /rpt\_family="MIR"  
5321. .5368  
repeat\_region /rpt\_family="L2"  
5447. .5471  
repeat\_region /rpt\_family="L1"  
5472. .5778  
repeat\_region /rpt\_family="Alu"  
5779. .6004  
repeat\_region /rpt\_family="L1"  
6005. .6317  
repeat\_region /rpt\_family="Alu"  
6318. .6674  
repeat\_region /rpt\_family="L1"  
6806. .7115  
repeat\_region /rpt\_family="Alu"  
7149. .7236  
repeat\_region /rpt\_family="L2"  
7501. .7634  
repeat\_region /rpt\_family="MIR"  
7932. .8041  
repeat\_region /rpt\_family="MIR"  
8242. .8538  
repeat\_region /rpt\_family="Alu"  
8546. .8667  
repeat\_region /rpt\_family="MIR"  
8804. .9106  
repeat\_region /rpt\_family="Alu"  
9627. .9661  
repeat\_region /rpt\_family="L1"  
10103. .10400  
repeat\_region /rpt\_family="Alu"  
10414. .10738  
repeat\_region /rpt\_family="L1"  
10739. .10971  
repeat\_region /rpt\_family="L1"  
11287. .11336  
repeat\_region /rpt\_family="MER1\_type"  
11342. .11473  
repeat\_region /rpt\_family="MER1\_type"  
11690. .11702  
repeat\_region /rpt\_family="Harlequin"  
11703. .11875  
repeat\_region /rpt\_family="(TTTC)n"  
11876. .11974  
repeat\_region /rpt\_family="Harlequin"

repeat\_region 11977. .12025  
/rpt\_family="Retroviral"  
repeat\_region 12027. .12295  
/rpt\_family="Alu"  
repeat\_region 12584. .12705  
/rpt\_family="MER1\_type"  
repeat\_region 12706. .12995  
/rpt\_family="Alu"  
repeat\_region 12996. .13026  
/rpt\_family="MER1\_type"  
repeat\_region 13027. .13171  
/rpt\_family="Alu"  
repeat\_region 13172. .13462

Query Match 1.7% Score 21; DB 86; Length 141899;  
Best Local Similarity 100.0%; Pred. No. 8.8; Indels 0; Gaps 0;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 322 attgcctgctcgaatcttt 342  
|||||  
Db 6441 ATTTGCTGCTCTGAATCTTT 6421

RESULT 8  
AC008627/c  
LOCUS AC008627 144793 bp DNA HTG 14-FEB-2001  
DEFINITION Homo sapiens chromosome 5 clone CTB-160022, WORKING DRAFT SEQUENCE,  
21 unordered pieces.  
ACCESSION AC008627  
VERSION AC008627.6 GI:12830083  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE human  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 144793)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 5  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 144793)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Feb 14, 2001 this sequence version replaced gi:9934573.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov  
-----  
Project Information  
Center Project Name: 126002, H278  
Center clone name: CIT978SKB-160Q22  
-----  
Summary Statistics  
Consensus quality: 122327 bases at least Q40  
Consensus quality: 134761 bases at least Q30  
Consensus quality: 137662 bases at least Q20  
Estimated insert size: 195000; pulse field gel estimation  
Estimated insert size: 142793; sum-of-contigs estimation  
Quality coverage: 7.8 in Q20 bases; pulse field gel estimation  
Quality coverage: 10.65 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1189: contig of 1189 bp in length  
\* 1190 1289: gap of unknown length  
\* 1290 2397: contig of 1108 bp in length



```

* 34050 34056: gap of unknown length
* 34057 52100: contig of 18044 bp in length
* 52101 52107: gap of unknown length
* 52108 61600: contig of 9493 bp in length
* 61601 61607: gap of unknown length
* 61608 77230: contig of 15623 bp in length
* 77231 77237: gap of unknown length
* 77238 79551: contig of 2314 bp in length
* 77238 79551: contig of 2314 bp in length
* 79552 79558: gap of unknown length
* 79559 83885: contig of 4327 bp in length.

FEATURES             Location/Qualifiers
     source            1..83885
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="16q24.3"
                        /clone="RP11-321A1"
                        /clone_lib="RP11"

BASE COUNT   19974 a 20796 c 20786 g 22201 t 128 others
ORIGIN

Query Match          1.7%: Score 21; DB 75; Length 83885;
Best Local Similarity 100.08; Pred. No. 8.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 ttggcgaagaaagaaataaca 85
|||||
Db 32237 TTTGGCAAAAGAAAAATACA 32257

RESULT 7
AC006334/c
LOCUS       AC006334      141899 bp      DNA      PRI      30-SEP-2000
DEFINITION Homo sapiens BAC clone RP11-421B22 from 7p11.2-q11.2, complete
            sequence.
ACCESSION   AC006334
VERSION     AC006334.3  GI:7622516
KEYWORDS    HTG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 141899)
AUTHORS     Sulston, J.E. and Waterston, R.
TITLE       Toward a complete human genome sequence
JOURNAL     Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE     99063792
REFERENCE   2 (bases 1 to 141899)
AUTHORS     Armstrong, J., Maupin, R. and Harris, A.
TITLE       The sequence of Homo sapiens BAC clone RP11-421B22
JOURNAL     Unpublished
REFERENCE   3 (bases 1 to 141899)
AUTHORS     Waterston, R.H.
TITLE       Direct Submission
JOURNAL     Submitted (09-JAN-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4 (bases 1 to 141899)
AUTHORS     Waterston, R.H.
TITLE       Direct Submission
JOURNAL     Submitted (20-APR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5 (bases 1 to 141899)
AUTHORS     Waterston, R.H.
TITLE       Direct Submission
JOURNAL     Submitted (21-APR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   6 (bases 1 to 141899)
AUTHORS     Waterston, R.H.
TITLE       Direct Submission
JOURNAL     Submitted (22-APR-2000) Genome Sequencing Center, Washington

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
-----
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
7 (bases 1 to 141899)
Waterston, R.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Apr 20, 2000 this sequence version replaced gi:6042131.
-----
Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
-----
Summary Statistics
-----
Center project name: H_NH0421B22
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/STB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RP11-421 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP5-845I21, 200 base pair overlap. Actual start of this clone is at base position 89122 of RP5-845I21; actual end is at base position 141899 of RP11-421B22.

#### FEATURES

```

source
1..141899
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7p11.2-q11.2"
/clone="RP11-421B22"
/clone_lib="RP11-11"
84..134
/rpt_family="Mariner"
135..433
/rpt_family="Alu"
434..485
/rpt_family="Mariner"
499..613
/rpt_family="MIR"
626..657
/rpt_family="Alu"
658..752
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

```

## submissions.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The sequence from clone L241B9 has been finished in more than one contig. This sequence (L241B9C) is separated from the preceding one (L241B9B) by a gap of about 100bp sized by restriction analysis. The true right end of clone L241B9 is at 17243.

L241B9 is from cosmid library LA04NC01 constructed at the Human Genome Center, Los Alamos National Laboratory, NM 87545 under the auspices of the U.S. Department of Energy. The library was constructed using flow-sorted human chromosome 4 from a Hamster-Human hybrid cell line (UV20HL21-27) containing human chromosomes 4, 8 and 21.

VECTOR: sCos1

L241B9 is contained in a clone contig spanning

2Mb which is described in Baxendale et al, Nature Genetics 4 ( 1993 ) 181- 186. See also Myers et al, Cytogenet Cell Genet. 66 ( 1994 ) 218-230.

## FEATURES

## source

Location/Qualifiers

```
1. .17243
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4p16.3"
/clone_lib="LA04NC01"
/cell_line="UV20HL21-27"
/clone="LA04NC01-241B9"
898. .1239
/partial
```

## misc\_feature

```
/note="match: VNTR M62748; M33007 locus PYN232;
Polymorphic compared with M62750; M33007; M33008"
complement(1569. .1728)
/partial
```

## misc\_feature

```
/note="match: EST L34968"
4342. .4415
```

## repeat\_region

```
/note="2 copies of 37 mer 89 & conserved"
6431. .6498
```

## repeat\_region

```
/note="4 copies of 17 mer 85 & conserved"
6458. .6539
```

## repeat\_region

```
/note="2 copies of 41 mer 90 & conserved"
7095. .7176
```

## repeat\_region

```
/note="2 copies of 41 mer 92 & conserved"
12869. .12903
```

## repeat\_region

```
/partial
/note="Alu repeat: matches 239. .205 of consensus"
12916. .12959
```

## repeat\_region

```
/partial
/note="Alu repeat: matches 44. .1 of consensus"
13018. .13435
```

## misc\_feature

```
/note="match: multiple ESTs"
13733. .13868
```

## repeat\_region

```
/partial
```

## repeat\_region

```
/note="Alu repeat: matches 1. .151 of consensus"
13869. .14160
```

## repeat\_region

```
/note="Alu repeat: matches 1. .308 of consensus"
14164. .14362
```

## repeat\_region

```
/partial
```

## repeat\_region

```
/note="Alu repeat: matches 122. .308 of consensus"
15471. .15765
```

## repeat\_region

```
/partial
```

## repeat\_region

```
/note="Alu repeat: matches 308. .1 of consensus"
15781. .16071
```

## repeat\_region

```
/partial
```

## repeat\_region

```
/note="Alu repeat: matches 308. .1 of consensus"
16332. .16630
```

## repeat\_region

```
/note="Alu repeat: matches 1. .308 of consensus"
5222 c 5198 g 3556 t
```

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## Query Match

1.7%; Score 21; DB 93; Length 17243;

## Local Similarity

100.0%; Pred. No. 8.8;

## Matches

21; Conservative 0; Mismatches 0; Gaps 0;

## Qy

37 ggagctggaagccatcaagt 57

## Db

15216 GGAGCTGGAGCCATCAAGTT 15236

## RESULT

AC074301

## LOCUS

AC074301 83895 bp DNA HTG 04-OCT-2000

## DEFINITION

Homo sapiens chromosome 16q24.3 clone RP11-321A1, WORKING DRAFT

## SEQUENCE

SEQUENCE, 18 ordered pieces.

## ACCESSION

AC074301

## VERSION

AC074301.2 GI:10280724

## KEYWORDS

HTG; HTGS\_PHASE2; HTGS\_DRAFT.

## SOURCE

human.

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 83895)

## AUTHORS

Kremmidiotis,G., Gardner,A.E., Callen,D.F. and Sutherland,G.R.

## JOURNAL

Large Scale Sequencing of the Chromosome 16 region q24.3

## REFERENCE

2 (bases 1 to 83895)

## AUTHORS

Kremmidiotis,G., Gardner,A.E., Callen,D.F. and Sutherland,G.R.

## JOURNAL

Submitted (26-JUL-2000) Cytogenetics & Molecular Genetics, Women's

## JOURNAL

& Children's Hospital / Bionomics, 28 Dalgleish Street Thebarton,

## JOURNAL

Adelaide, South Australia 5031, Australia

## JOURNAL

On Sep 23, 2000 this sequence version replaced gi:9454506.

## COMMENT

Genome Centre : Cytogenetics & Molecular Genetics

Centre Code : CMGWCH

Website: http://www.wch.sa.gov.au/labmedic/genetics/sequencing.html

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 18 contigs. Gaps between the contigs

\* are represented as runs of N. The order of the pieces

\* is believed to be correct as given, however the sizes

\* of the gaps between them are based on estimates that have

\* provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 7565: contig of 7565 bp in length

\* 7566 7573: gap of unknown length

\* 7574 10623: contig of 3050 bp in length

\* 10624 10631: gap of unknown length

\* 10632 16358: contig of 5727 bp in length

\* 16359 16366: gap of unknown length

\* 16367 28854: contig of 12488 bp in length

\* 28855 28861: gap of unknown length

\* 28862 30975: contig of 2114 bp in length

\* 30976 30982: gap of unknown length

\* 30983 31699: contig of 717 bp in length

\* 31700 31706: gap of unknown length

\* 31707 32399: contig of 693 bp in length

\* 32400 32406: gap of unknown length

\* 32407 32716: contig of 310 bp in length

\* 32717 32723: gap of unknown length

\* 32724 33243: contig of 520 bp in length

\* 33244 33250: gap of unknown length

\* 33251 33413: contig of 163 bp in length

\* 33414 33420: gap of unknown length

\* 33421 33540: contig of 120 bp in length

\* 33541 33547: gap of unknown length

\* 33548 33996: contig of 449 bp in length

\* 33997 34003: gap of unknown length

\* 34004 34049: contig of 46 bp in length

```

repeat_region 29832..30597
/note="L1MC1 repeat: matches 4945..5745 of consensus"
repeat_region 30598..30732
/note="MER85 repeat: matches 2..138 of consensus"
repeat_region 30733..31319
/note="L1MC1 repeat: matches 5745..6320 of consensus"
repeat_region 31344..31944
/note="MER67C repeat: matches 74..710 of consensus"
repeat_region 32736..33172
/note="L2 repeat: matches 2252..2719 of consensus"
repeat_region 33142..33526
/note="L2 repeat: matches 1853..2276 of consensus"
repeat_region 33758..34056
/note="L2 repeat: matches 2102..2418 of consensus"
repeat_region 36152..36234
/note="L2 repeat: matches 2641..2738 of consensus"
/note="match: GSS: Em:AQ784289"
misc_feature 37890..38425
/note="match: GSS: Em:AQ561026"
repeat_region 38199..38409
/note="MER58A repeat: matches 2..224 of consensus"
repeat_region 38825..38875
/note="HERVL repeat: matches 4905..4956 of consensus"
repeat_region 39393..39717
/note="L1R16A repeat: matches 96..437 of consensus"
repeat_region 40729..40876
/note="MIR repeat: matches 105..262 of consensus"
repeat_region 41852..42041
/note="ML1A1 repeat: matches 189..365 of consensus"
repeat_region 42042..42345
/note="AluY repeat: matches 1..304 of consensus"
repeat_region 42346..42495

Query Match 1.8%; Score 22; DB 90; Length 129654;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1072 tattaaagcacatgtttgaaa 1093
|||||
Db 10095 TATTAAGACATGTTTGA 10116

RESULT 4
AL355516 158427 bp DNA HTG 08-APR-2001
LOCUS Homo sapiens chromosome 13 clone RP11-442J17, *** SEQUENCING IN
DEFINITION PROGRESS ***, 3 unordered pieces.
ACCESSION AL355516
VERSION AL355516.10 GI:13568039
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 158427)
Tracey,A.
Direct Submission
Submitted (07-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Apr 9, 2001 this sequence version replaced gi:13373936.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA442J17
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 99% of reads

```

```

Chemistry: Dye-primer Big Dye; 0% of reads
Consensus quality: 156821 bases at least Q40
Consensus quality: 157434 bases at least Q30
Consensus quality: 157904 bases at least Q20
Insert size: 158227; sum-of-contigs
Insert size: 162425; 5.4% error; agarose-fp
Quality coverage: 11.86x in Q20 bases; sum-of-contigs Quality
coverage: 12.36x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 154082: contig of 154082 bp in length
* 154083 154182: gap of 100 bp
* 154183 156245: contig of 2063 bp in length
* 156246 156345: gap of 100 bp
* 156346 158427: contig of 2082 bp in length.

FEATURES
Location/Qualifiers
source 1..158427
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-442J17"
/clone_lib="RPC1-11.2"
misc_feature 1..154082
/note="assembly_fragment:02255
clone_end:SP6
vector_side:left
clone_end:T7
vector_side:right"
misc_feature 154183..156245
/note="assembly_fragment:00968"
misc_feature 156346..158427
/note="assembly_fragment:04620"
BASE COUNT 47544 a 34501 c 32262 g 43917 t 203 others
ORIGIN

Query Match 1.8%; Score 22; DB 79; Length 158427;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1072 tattaaagcacatgtttgaaa 1093
|||||
Db 156460 TATTAAGACATGTTTGA 156481

RESULT 5
HSL241B9C 17243 bp DNA PRI 23-NOV-1999
LOCUS Human DNA sequence from cosmid L241B9, Huntington's Disease Region,
DEFINITION Chromosome 4p16.3 contains polymorphic VNTR pYN232.
ACCESSION Z69708
VERSION Z69708.1 GI:1204106
KEYWORDS 4p16.3; repeat polymorphism.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 17243)
Mungall,A and Odell,C.
Direct Submission
Submitted (13-FEB-1996) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, UK. E-mail enquiries: humquery@sanger.ac.uk
IMPORTANT:
This sequence is not the entire insert of clone L241B9. It may be
shorter because we only sequence overlapping sections once, or
longer because we arrange for a small overlap between neighbouring

```

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13>

IMPORTANT: This sequence is not the entire insert of clone RP11-408L13 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-408L13 is at 1 in this sequence. The true left end of clone RP11-2140L1 is at 129555 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-408L13 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.choiri.org/bacpac/home.htm>

#### FEATURES

```

source
    Location/Qualifiers
        1..129654
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="13"
            /clone="RP11-408L13"
            /clone_lib="RPCI-11.2"
repeat_region
    1..239
        /note="L2 repeat: matches 1911..2196 of consensus"
misc_feature
    12..525
        /note="match: GSS: Em:AQ336465"
misc_feature
    330..759
        /note="match: STS: Em:HSPH27F11"
repeat_region
    643..864
        /note="111 copies 2 mer aa 55% conserved"
repeat_region
    873..927
        /note="FLAM repeat: matches 87..133 of consensus"
repeat_region
    1065..1297
        /note="L2 repeat: matches 5567..5780 of consensus"
repeat_region
    1426..1465
        /note="L2 repeat: matches 2695..2734 of consensus"
misc_feature
    1509..2013
        /note="match: GSS: Em:AQ370915"
misc_feature
    1518..1925
        /note="match: GSS: Em:AQ450615"
repeat_region
    1532..1597
        /note="38 copies 2 mer ta 68% conserved"
repeat_region
    1819..1888
        /note="35 copies 2 mer aa 65% conserved"
repeat_region
    2488..2558
        /note="MIR repeat: matches 84..154 of consensus"
repeat_region
    3285..3388
        /note="Charlie5 repeat: matches 2479..2582 of consensus"
repeat_region
    3391..3576
        /note="match: GSS: Em:AQ268552"
repeat_region
    3602..3742
        /note="MER8 repeat: matches 54..239 of consensus"
misc_feature
    complement(3937..4380)
        /note="match: GSS: Em:AQ268552"
misc_feature
    complement(3994..4383)
        /note="match: GSS: Em:AQ518593"
repeat_region
    4242..4623
        /note="L1M3c repeat: matches 10..401 of consensus"
repeat_region
    5230..5414
        /note="L1M1 repeat: matches 1880..2452 of consensus"
repeat_region
    5427..6545

```

```

/note="L1MEC repeat: matches 780..1969 of consensus"
6589..6753
/note="L1ME2 repeat: matches 5945..6100 of consensus"
6872..7011
/note="MIR repeat: matches 4..146 of consensus"
8145..8514
/note="MLR1D repeat: matches 29..433 of consensus"
9369..9537
/note="MER20 repeat: matches 30..218 of consensus"
9554..9661
/note="54 copies 2 mer at 66% conserved"
10799..11307
/note="match: GSS: Em:AQ572234"
11782..12049
/note="AlusX repeat: matches 31..297 of consensus"
12051..12473
/note="L1R12 repeat: matches 3..437 of consensus"
12478..13246
/note="HERV9 repeat: matches 7631..8399 of consensus"
13247..13280
/note="AlusX repeat: matches 1..34 of consensus"
13859..14126
/note="AluJ0 repeat: matches 1..269 of consensus"
14195..14576
/note="L2 repeat: matches 1875..2264 of consensus"
15738..16230
/note="MLR1D repeat: matches 1..505 of consensus"
complement(18427..18872)
/note="match: GSS: Em:AQ137188"
19011..19400
/note="L1M2 repeat: matches 260..656 of consensus"
19404..19633
/note="L1M2 repeat: matches 830..1064 of consensus"
19634..19939
/note="AlusX repeat: matches 1..307 of consensus"
19940..19975
/note="L1M2 repeat: matches 1064..1099 of consensus"
19976..20245
/note="AluJ0 repeat: matches 5..290 of consensus"
20246..20615
/note="L1M2 repeat: matches 1099..1582 of consensus"
20863..20976
/note="L1M2 repeat: matches 1582..1697 of consensus"
20988..21169
/note="L1M2 repeat: matches 1705..1883 of consensus"
21297..21573
/note="AlusX repeat: matches 20..293 of consensus"
complement(21311..21860)
/note="match: GSS: Em:AQ453572"
21709..21882
/note="L1PBA repeat: matches -1552..-1382 of consensus"
21884..22388
/note="L1P3 repeat: matches 5640..6146 of consensus"
22419..22724
/note="AlusG repeat: matches 1..310 of consensus"
22725..23155
/note="L1PBA repeat: matches -1393..-964 of consensus"
23162..23918
/note="L1PBA repeat: matches -260..-1443 of consensus"
23925..24543
/note="L1 repeat: matches 2420..3027 of consensus"
24544..24859
/note="L1P5 repeat: matches 5393..5746 of consensus"
24845..26144
/note="L1P repeat: matches 1924..3208 of consensus"
26137..27882
/note="L1P7 repeat: matches 3779..5587 of consensus"
27883..28325
/note="L1P7 repeat: matches 5698..6143 of consensus"
28351..28791
/note="L1P2 repeat: matches 5737..6150 of consensus"
28857..29817
/note="L1MEC repeat: matches 2390..3028 of consensus"

```

repeat\_region /note="AluJb repeat: matches 7. .298 of consensus"  
19568. .20654  
repeat\_region /note="L1ME1 repeat: matches 4980. .6159 of consensus"  
20853. .21023  
repeat\_region /note="L1M1 repeat: matches 5362. .5534 of consensus"  
22340. .22924  
repeat\_region /note="L1MEC repeat: matches 1138. .1691 of consensus"  
22945. .23582  
repeat\_region /note="MER82 repeat: matches 11. .652 of consensus"  
23834. .23939  
repeat\_region /note="53 copies 2 mer aa 60% conserved"  
23955. .24313  
repeat\_region /note="L1M4 repeat: matches 2122. .2530 of consensus"  
24537. .24842  
repeat\_region /note="AluJ repeat: matches 1. .305 of consensus"  
25067. .25266  
repeat\_region /note="MLA1E repeat: matches 186. .391 of consensus"  
25585. .26159  
repeat\_region /note="L1MD2 repeat: matches 3360. .3993 of consensus"  
26276. .27473  
repeat\_region /note="L1MD2 repeat: matches 4146. .5398 of consensus"  
27602. .28031  
repeat\_region /note="L1MD2 repeat: matches 5470. .5893 of consensus"  
28149. .28245  
repeat\_region /note="L2 repeat: matches 2576. .2678 of consensus"  
29003. .29160  
repeat\_region /note="MIR repeat: matches 67. .228 of consensus"  
30465. .30965  
misc\_feature /note="match: GSS: Em:AQ702118"  
30730. .30805  
repeat\_region /note="L2 repeat: matches 2635. .2710 of consensus"  
31335. .31591  
repeat\_region /note="L1M4 repeat: matches 4776. .5036 of consensus"  
31756. .33219  
repeat\_region /note="L1P81 repeat: matches 4677. .6148 of consensus"  
33278. .33614  
repeat\_region /note="L1WC/D repeat: matches 5263. .5589 of consensus"  
33623. .33965  
repeat\_region /note="MLR1H repeat: matches 5. .401 of consensus"  
34146. .34503  
repeat\_region /note="THE1B repeat: matches 1. .364 of consensus"  
34958. .35177  
repeat\_region /note="MIR repeat: matches 12. .235 of consensus"  
35195. .35230  
repeat\_region /note="9 copies 4 mer acac 100% conserved"  
35306. .35601  
repeat\_region /note="AluJb repeat: matches 2. .299 of consensus"  
36464. .36737  
repeat\_region /note="AluJb repeat: matches 39. .301 of consensus"  
complement(37097. .37453)  
misc\_feature /note="match: GSS: Em:AQ527226"  
38095. .38735  
repeat\_region /note="L2 repeat: matches 1783. .2455 of consensus"  
39500. .39535  
repeat\_region /note="18 copies 2 mer aa 91% conserved"  
39998. .40216  
repeat\_region /note="MIR repeat: matches 21. .262 of consensus"  
40356. .40495  
repeat\_region /note="L2 repeat: matches 1198. .1357 of consensus"  
41485. .41687  
repeat\_region /note="L2 repeat: matches 2050. .2266 of consensus"  
41763. .41914  
repeat\_region /note="AluJp/q repeat: matches 160. .313 of consensus"  
41915. .41992  
repeat\_region /note="3 copies 26 mer 79% conserved"  
42256. .42707  
repeat\_region /note="L2 repeat: matches 862. .1277 of consensus"  
42976. .43601  
misc\_feature /note="match: GSS: Em:AQ554371"  
43347. .43651  
repeat\_region /note="AluJ repeat: matches 1. .303 of consensus"  
44738. .45083  
repeat\_region /note="L2 repeat: matches 274. .635 of consensus"

45269. .45460  
repeat\_region /note="MIR repeat: matches 21. .235 of consensus"  
48328. .48496  
repeat\_region /note="MER5A repeat: matches 18. .189 of consensus"  
48616. .48707  
repeat\_region /note="MER5A repeat: matches 63. .151 of consensus"  
49220. .49493  
repeat\_region /note="L2 repeat: matches 1667. .1942 of consensus"  
complement(50916. .51036)  
misc\_feature /note="match: GSS: Em:B42447"  
51876. .52224  
repeat\_region /note="L2 repeat: matches 2369. .2705 of consensus"  
52897. .53334  
misc\_feature /note="match: GSS: Em:AQ267168"  
53551. .53634  
repeat\_region /note="MIR repeat: matches 45. .132 of consensus"  
54751. .54843  
repeat\_region /note="L2 repeat: matches 2548. .2648 of consensus"  
55244. .55500  
repeat\_region /note="L2 repeat: matches 2442. .2746 of consensus"  
complement(56461. .56862)  
misc\_feature /note="match: GSS: Em:AQ196491"  
complement(56570. .56862)  
misc\_feature /note="match: GSS: Em:B92720"  
57784. .57842  
repeat\_region /note="MIR repeat: matches 50. .112 of consensus"  
complement(57842. .58272)  
misc\_feature /note="match: GSS: Em:B49733"  
Query Match 16.0%; Score 194; DB 90; Length 106585;  
Best Local Similarity 100.0%; Pred. NO. 3.1e-93;  
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 799 agttccatcaagctccgctcagagaatgctatgttgatccatataaaactcacatttc 858  
|||||  
Db 52944 AGTTCATCAAGCTCCGCTCAGAGAATGCTATGGTGATCCATATAAAATCATTTC 52885  
Qy 859 cagtgcagcgtgtctctgaattcctgtcactgtcagatattattctaaagtcacccc 918  
|||||  
Db 52884 CAGTGCAGCGCTGCTCTGAATTCCTGCATACGTGCAGATATTATTCTAAAGTCACCCC 52825  
Qy 919 ttctagatcaactgctgtaacaaatcagaacatttcaatttcatttcacgttcagcat 978  
|||||  
Db 52824 TTCCTACCATGCAACTGCTGTAAACAATCAGAACATTTCAATTTACCTTCAGCCAT 52765  
Qy 979 gcccccggtctcaa 992  
|||||  
Db 52764 GCCCCCGGTCTCAA 52751  
RESULT 3  
AL157765  
LOCUS AL157765 129654 bp DNA PRI 26-MAR-2001  
DEFINITION Human DNA sequence from clone RP11-408L13 on chromosome 13 Contains  
STSS and GSSs, complete sequence.  
ACCESSION AL157765  
VERSION AL157765.6. GI:11228550  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 129654)  
AUTHORS Kay, M.  
TITLE Direct Submission  
JOURNAL Submitted (16-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
COMMENT On Nov 20, 2000 this sequence version replaced gi:11125411.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2001, 10:00:07 ; Search time 2169.51 seconds  
(without alignments)  
8662.470 Million cell updates/sec

Title: US-09-389-000-1\_copy\_735\_1949  
Perfect score: 1215  
Sequence: 1 ggtctctgaagtgccctt.....tiggcatggttggctctg 1215

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0  
Searched: 1344157 seqs, 7733874588 residues  
Word size : 0

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba1.\*
- 2: gb\_ba2.\*
- 3: gb\_ba3.\*
- 4: gb\_in1.\*
- 5: gb\_in2.\*
- 6: gb\_in3.\*
- 7: gb\_om.\*
- 8: gb\_ov.\*
- 9: gb\_pat1.\*
- 10: gb\_pat2.\*
- 11: gb\_ph.\*
- 12: gb\_pl1.\*
- 13: gb\_pl2.\*
- 14: gb\_pl3.\*
- 15: gb\_pl4.\*
- 16: gb\_ba1.\*
- 17: em\_ba2.\*
- 18: em\_fun.\*
- 19: em\_htgo\_hum.\*
- 20: em\_htgo\_inv.\*
- 21: em\_htgo\_rod.\*
- 22: em\_htg\_hum1.\*
- 23: em\_htg\_hum2.\*
- 24: em\_htg\_hum3.\*
- 25: em\_htg\_hum4.\*
- 26: em\_htg\_hum5.\*
- 27: em\_htg\_hum6.\*
- 28: em\_htg\_hum7.\*
- 29: em\_htg\_hum8.\*
- 30: em\_htg\_inv1.\*
- 31: em\_htg\_inv2.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_rod.\*
- 34: em\_hum1.\*
- 35: em\_hum2.\*
- 36: em\_hum3.\*
- 37: em\_hum4.\*
- 38: em\_hum5.\*
- 39: em\_hum6.\*
- 40: em\_hum7.\*
- 41: em\_in.\*
- 42: em\_om.\*
- 43: em\_or.\*

- 44: em\_ov.\*
- 45: em\_pat.\*
- 46: em\_ph.\*
- 47: em\_pl.\*
- 48: em\_ro.\*
- 49: em\_sts.\*
- 50: em\_sy.\*
- 51: em\_un.\*
- 52: em\_vi.\*
- 53: gb\_sts1.\*
- 54: gb\_sts2.\*
- 55: gb\_sts3.\*
- 56: gb\_sy.\*
- 57: gb\_un.\*
- 58: gb\_vil.\*
- 59: gb\_vil2.\*
- 60: gb\_htg1.\*
- 61: gb\_htg2.\*
- 62: gb\_htg3.\*
- 63: gb\_htg4.\*
- 64: gb\_htg5.\*
- 65: gb\_htg6.\*
- 66: gb\_htg7.\*
- 67: gb\_htg8.\*
- 68: gb\_htg9.\*
- 69: gb\_htg10.\*
- 70: gb\_htg11.\*
- 71: gb\_htg12.\*
- 72: gb\_htg13.\*
- 73: gb\_htg14.\*
- 74: gb\_htg15.\*
- 75: gb\_htg16.\*
- 76: gb\_htg17.\*
- 77: gb\_htg18.\*
- 78: gb\_htg19.\*
- 79: gb\_htg20.\*
- 80: gb\_htg21.\*
- 81: gb\_htg22.\*
- 82: gb\_htg23.\*
- 83: gb\_htg24.\*
- 84: gb\_htg25.\*
- 85: gb\_pr1.\*
- 86: gb\_pr2.\*
- 87: gb\_pr3.\*
- 88: gb\_pr4.\*
- 89: gb\_pr5.\*
- 90: gb\_pr6.\*
- 91: gb\_pr7.\*
- 92: gb\_pr8.\*
- 93: gb\_pr9.\*
- 94: gb\_rol.\*
- 95: gb\_rol2.\*
- 96: gb\_in4.\*
- 97: gb\_pr10.\*
- 98: em\_ba3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	942	77.5	2147	89	AK000456 Homo sapi
2	194	16.0	106585	90	AL160392 Human DNA
3	22	1.8	129654	90	AL157765 Human DNA
4	22	1.8	158427	79	AL355516 Homo sapi
5	21	1.7	17243	93	HSL241B9C
6	21	1.7	83885	75	AC074301 Homo sapi
7	21	1.7	141899	86	AC006334 Homo sapi
8	21	1.7	144793	60	AC008627 Homo sapi





```

/lab_host="DH12S"
/Note="Organ: whole embryo; Vector: pCMV-SPORT; Site_1:
Salt; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Gastrulating embryos were collected at 7.5dpc
from C57BL6 x DBA matings, excluding embryos that had
developed head folds and all extraembryonic tissues.
Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
Referenced in Development 121, 2479-2489 (1995)"
BASE COUNT      59 a   47 c   46 g      87 t
ORIGIN

```

```

Query Match      1.6%; Score 20; DB 3; Length 249;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 72 aaagagaaatacacattca 91
      |||||
Db 191 AAAAGAAAATACACATTCA 172

```

```

RESULT 15
BE606386/c
LOCUS      BE606386      319 bp      mRNA      EST      22-AUG-2000
DEFINITION WHE0904_D06_H12S Wheat 5-15 DAP spike cDNA library Triticum
aestivum cDNA clone WHE0904_D06_H12, mRNA sequence.
ACCESSION BE606386
VERSION BE606386.1 GI:9883550
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
; Triticeae; Triticum.
1 (bases 1 to 319)
Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat
genomes - 5-15 DAP spike cDNA library
Unpublished (2000)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

```

```

FEATURES
source
1..319
location/Qualifiers
    /organism="Triticum aestivum"
    /cultivar="Chinese Spring"
    /db_xref="taxon:4565"
    /clone="WHE0904_D06_H12"
    /clone_lib="Wheat 5-15 DAP spike cDNA library"
    /tissue_type="Spike"
    /dev_stage="Adult plant"
    /lab_host="E. coli SOLR"
    /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
greenhouse. Spikes at 5, 10 and 15 DAP were harvested,
total RNA and poly(A) RNA were prepared, a cDNA library
was made, and the cDNA clones were in vivo excised to
give pBluescript phagemids in the TJ Close lab (Choi,
Close, Fenton) at the University of California,
Riverside. Plasmid DNA preparations and DNA sequencing
were performed in the OD Anderson lab (all other authors
)."
```

```

BASE COUNT      78 a   86 c   108 g   47 t
ORIGIN

```

```

Query Match      1.6%; Score 20; DB 137; Length 319;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 981 cccgggtctcaagcttctcc 1000
      |||||
Db 111 CCCGGTCTCAAGCTTCTCC 92

```

```

Search completed: September 15, 2001, 12:34:24
Job time: 5874 sec

```

JOURNAL  
COMMENT Unpublished (2000)  
Contact: Cathy Konning  
The Institute for Genomic Research  
For clone info: please contact Research Genetics, Libraries  
Division tel 1-800-711-6195, email cdna@resgen.com  
Seq primer: M13F-R.

FEATURES  
Source Location/Qualifiers  
1..606  
/organism="Solanum tuberosum"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="BPLI2J13"  
/clone\_lib="P.infestans-challenged leaf"  
/tissue\_type="leaf"  
/dev\_stage="6 week old"  
/lab\_host="SOLR"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; Whole plants were challenged with 450,000  
sporangia/ml P. infestans US-1(US 940501) in Biotron  
(Madison, Wisconsin). Leaf tissue was collected at 1, 2,  
5, 12, and 24 hours post-challenge and frozen in liquid  
nitrogen immediately upon removal. Kennebec plants showed  
no signs of HR. Katahdin plants (susceptible to P.  
infestans US-1) were used as controls and showed  
infection. NOTE: We cannot exclude the possibility that  
this sequence is actually derived from Phytophthora rather  
than potato."

BASE COUNT 179 a 100 c 141 g 186 t  
ORIGIN

Query Match 1.7%; Score 21; DB 155; Length 606;  
Best Local Similarity 100.0%; Pred. No. 6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 agatctctctcttcattcaa 422  
|||||  
Db 4 AGATCTCTCTTCTTCATCAA 24

RESULT 13  
LOCUS AQ659253/625 bp DNA GSS 23-JUN-1999  
DEFINITION Sheared DNA-3D4.TF Sheared DNA Trypanosoma brucei genomic clone  
Sheared DNA-3D4, DNA sequence.  
ACCESSION AQ659253  
VERSION AQ659253.1 GI:5167021  
KEYWORDS GSS  
SOURCE Trypanosoma brucei.  
ORGANISM Trypanosoma brucei  
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;  
Trypanosoma.  
1 (bases 1 to 625)  
El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,  
Gerrard,C., Leech,N., de Jong,P., Ollu,E., Melville,S., Doneison,J.,  
Fraser,C. and Adams,M.  
Determination of clone end sequences from Trypanosoma brucei GUTat  
10.1 sheared DNA library  
Unpublished (1999)  
Other GSSs: Sheared DNA-3D4.TF  
Contact: Najib M. El-Sayed  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: nelsayed@tigr.org  
Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared  
DNA library constructed at TIGR. Clones will be available for  
distribution through AFCC. Sheared DNA end sequences search page:  
<http://www.tigr.org/tldb/mdb/tbdb/>.  
Seq primer: M13-Reverse  
Class: Shotgun.

FEATURES  
Source Location/Qualifiers  
1..625  
/organism="Trypanosoma brucei"  
/strain="TREU927/4 GUTat 10.1"  
/db\_xref="taxon:5691"  
/clone="Sheared DNA-3D4"  
/clone\_lib="Sheared DNA"  
/note="Vector: pUC18; Site\_1: SmaI; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from a cloned population of  
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically  
sheared to give a tight size distribution (approx 2 kb).  
The v + i method used for the library construction is  
described in detail in Smith, H.O. and Venter, J.C.  
(Making small insert libraries for whole genome shotgun  
sequencing projects. In Genome Sequencing: A Practical  
Approach, eds. M. Vaudin and B. Borell, Oxford University  
Press, 1999)."

BASE COUNT 174 a 155 c 154 g 142 t  
ORIGIN

Query Match 1.7%; Score 21; DB 231; Length 625;  
Best Local Similarity 100.0%; Pred. No. 6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 488 tatgtataaaggagagaagaat 508  
|||||  
Db 115 TATGTATAAAGGAGAGAAGAAT 95

RESULT 14  
LOCUS AAL155374/249 bp mRNA EST 11-DEC-1996  
DEFINITION mn43h12.rl beddington mouse embryonic region Mus musculus CDNA  
clone IMAGE:540743 5', mRNA sequence.  
ACCESSION AAL155374  
VERSION AAL155374.1 GI:1727004  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 249)  
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
The WashU-HMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of MedicineP  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:327679  
Seq primer: -40m13 ET  
High quality sequence stop: 236.  
Location/Qualifiers  
1..249  
/organism="Mus musculus"  
/strain="C57BL6 x DBA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:540743"  
/clone\_lib="Beddington mouse embryonic region"  
/sex="pooled"  
/tissue\_type="embryo"  
/dev\_stage="7.5dpc"

BASE COUNT 45 a 49 c 5 g 71 t  
ORIGIN

stages of tuber formation."

Query Match 1.7% Score 21; DB 166; Length 170;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 agatctctcttcttcatcaa 422

|||||

Db 109 AGATCTCTCTTCTTCATCAA 129

RESULT 10  
BF187644

LOCUS BF187644 471 bp mRNA EST 01-NOV-2000  
DEFINITION EST443931 potato stolon, Cornell University Solanum tuberosum cDNA  
clone CSTA41B22 5' sequence, mRNA sequence.

ACCESSION BF187644

VERSION BF187644.1 GI:11069863

KEYWORDS EST.

SOURCE potato.

ORGANISM Solanum tuberosum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE 1

AUTHORS van der Hoeven, R., Bezzerides, J., Bachem, C., Horvath, B., Visser, R.,  
Holt, I.E., Liang, F., Hansen, T.S., Utterback, T., Bowman, C.L., Doan  
, B., Bougri, O., Buell, C.R., Ronning, C.M., Tanksley, S.D. and Baker

TITLE Generation of ESTs from potato swelling stolons

JOURNAL Unpublished (1999)

COMMENT Contact: TIGR

The Institute for Genomic Research

Rockville, MD 20850, USA

Tel: (301)838-0200

Email: <http://www.tigr.org/tdb/potato/>

For clone information, please contact Research Genetics, Libraries  
Division, tel# 1-800-711-6195, email- [cdna@resgen.com](mailto:cdna@resgen.com).

FEATURES

source

1. .471

/organism="Solanum tuberosum"

/cultivar="Bintje"

/db\_xref="taxon:4113"

/clone="CSTA41B22"

/clone\_lib="potato stolon, Cornell University"

/tissue\_type="axillary buds of stem explants, swelling

stolons"

/dev\_stage="1 to 3 days"

/lab\_host="SOLR"

/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:

XhoI; RNA was supplied by Christian Bachem & Beatrix

Horvath(Laboratory of Plant Breeding, Dept. of Plant

Sciences, Wageningen University, The Netherlands). Total

RNA was isolated from developing axillary buds of potato

nodal stem cuttings cultured on medium for the

introduction of tuber formation as described in Bachem et

al. (Plant Journal 1996). Tissue samples were taken of

stages corresponding to growing stolons and the early

stages of tuber formation."

125 a 110 c 68 g 168 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 1.7% Score 21; DB 145; Length 471;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 agatctcttcttcatcaa 422

|||||

Db 175 AGATCTCTCTTCTTCATCAA 195

RESULT 11

TA268D06P

LOCUS

DEFINITION

genomic survey sequence.

ACCESSION

AL484903

VERSION

AL484903.1

KEYWORDS

GSS.

ORGANISM

Trypanosoma brucei.

Trypanosoma brucei

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma

1 (bases 1 to 560)

REFERENCE

AUTHORS

Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,

Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,

Melville, S.E., Rajandream, M.A. and Barrell, B.G.

Direct Submission

Submitted (10-DEC-2000)

project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,

Cambridge CB10 1SA, E-mail: [barrell@sanger.ac.uk](mailto:barrell@sanger.ac.uk) and

[nhl@sanger.ac.uk](mailto:nhl@sanger.ac.uk)

COMMENT

Constructed at the Institute for Genomic Research (TIGR),

Rockville, MD. Genomic DNA isolated from a cloned population of

Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared

to give a tight size distribution (

4 kb). The v + i method used for the library construction is

described in detail in Smith, H. and Venter, J.C. (Making small

insert libraries for whole genome shotgun sequencing projects. In

Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barrell, Oxford University Press, 1999).

Email: [nelsayed@tigr.org](mailto:nelsayed@tigr.org)

Details of T. brucei sequencing at the Sanger Centre are available

at [http://www.sanger.ac.uk/projects/T\\_brucei/](http://www.sanger.ac.uk/projects/T_brucei/).

FEATURES

Location/Qualifiers

1. .560

/organism="Trypanosoma brucei"

/strain="TREU927"

/db\_xref="taxon:5691"

/clone="268D06"

BASE COUNT 155 a 128 c 149 g 128 t

ORIGIN

Qy 488 tatgtataaaggagagaagaat 508

|||||

Db 133 TATGTAAAGGGAGAGAAAGAT 153

RESULT 12

BG589678

LOCUS

DEFINITION

EST497520 P. infestans-challenged leaf Solanum tuberosum cDNA clone

BDU2J13 5' sequence, mRNA sequence.

ACCESSION BG589678

VERSION BG589678.1

KEYWORDS EST.

SOURCE potato.

ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 606)

AUTHORS Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukolanov, A.,

Rangel, P., Haberlach, G.T., Cho, J., Chiemingo, A., Bougri, O., Buell

, C.R., Ronning, C.M., Helgeson, J. and Baker, B.

Generation of ESTs from Potato Leaves Challenged with Phytophthora

infestans, Incompatible Reaction

TITLE

```

and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGACG
Plate: 40 row: L column: 7
Seq primer: ATTAGGTGACACATATAG.
Location/Qualifiers
1. 553
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 163 a 106 c 121 g 163 t
ORIGIN

Query Match 2.1%; Score 26; DB 144; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 aattttaaaaggttgtatttcaggcca 141
|||||
Db 247 AATTTAAAGGTTGTATTTCAGGGCA 272

RESULT 9
BE342634 170 bp mRNA EST 17-JUL-2000
LOCUS
DEFINITION EST395478 potato stolon, Cornell University Solanum tuberosum cDNA
clone cSTA20P1, mRNA sequence.
ACCESSION BE342634
VERSION BE342634.1 GI:9252166
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 170)
AUTHORS van der Hoeven,R., Bezzerides,J., Bachem,C., Horvath,B., Visser,R.,
Holt,I.E., Liang,F., Hansen,T.S., Utterback,T., Bowman,C.L., Doan
,B., Bougri,O., Buell,C.R., Ronning,C.M., Tanksley,S.D. and Baker
,B.
TITLE Generation of ESTs from potato swelling stolons
JOURNAL Unpublished (1999)
COMMENT Contact: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cdna@resgen.com
5 prime sequence.
Location/Qualifiers
1. 170
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="cSTA20P1"
/clone_lib="potato stolon, Cornell University"
/tissue_type="axillary buds of stem explants, swelling
stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; RNA was supplied by Christian Bachem & Beatrix
Horvath(Laboratory of Plant Breeding, Dept. of Plant
Sciences, Wageningen University, The Netherlands). Total
RNA was isolated from developing axillary buds of potato
nodal stem cuttings cultured on medium for the
introduction of tuber formation as described in Bachem et
al. (Plant Journal 1996). Tissue samples were taken of
stages corresponding to growing stolons and the early

```

---

```

FEATURES
source
1. 624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGF"
/note="vector: pBluescriptSKm"
BASE COUNT 209 a 117 c 128 g 170 t
ORIGIN

Query Match 3.5%; Score 43; DB 122; Length 624;
Best Local Similarity 100.0%; Pred. No. 5.1e-12;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 ttactaagtcgagagcgtggaagccatcaagttaattaga 64
|||||
Db 462 TTCCTAAGTCCGAGGAGCGTGGAGCCATCAAGTTAATAGA 504

RESULT 8
BF080785 553 bp mRNA EST 18-OCT-2000
LOCUS
DEFINITION 231799 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF080785
VERSION BF080785.1 GI:10874615
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 553)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18

```

---

```

FEATURES
source
1. 624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGF"
/note="vector: pBluescriptSKm"
BASE COUNT 209 a 117 c 128 g 170 t
ORIGIN

Query Match 3.5%; Score 43; DB 122; Length 624;
Best Local Similarity 100.0%; Pred. No. 5.1e-12;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 ttactaagtcgagagcgtggaagccatcaagttaattaga 64
|||||
Db 462 TTCCTAAGTCCGAGGAGCGTGGAGCCATCAAGTTAATAGA 504

RESULT 8
BF080785 553 bp mRNA EST 18-OCT-2000
LOCUS
DEFINITION 231799 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF080785
VERSION BF080785.1 GI:10874615
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 553)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18

```

---

```

FEATURES
source
1. 553
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 163 a 106 c 121 g 163 t
ORIGIN

Query Match 2.1%; Score 26; DB 144; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 aattttaaaaggttgtatttcaggcca 141
|||||
Db 247 AATTTAAAGGTTGTATTTCAGGGCA 272

RESULT 9
BE342634 170 bp mRNA EST 17-JUL-2000
LOCUS
DEFINITION EST395478 potato stolon, Cornell University Solanum tuberosum cDNA
clone cSTA20P1, mRNA sequence.
ACCESSION BE342634
VERSION BE342634.1 GI:9252166
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 170)
AUTHORS van der Hoeven,R., Bezzerides,J., Bachem,C., Horvath,B., Visser,R.,
Holt,I.E., Liang,F., Hansen,T.S., Utterback,T., Bowman,C.L., Doan
,B., Bougri,O., Buell,C.R., Ronning,C.M., Tanksley,S.D. and Baker
,B.
TITLE Generation of ESTs from potato swelling stolons
JOURNAL Unpublished (1999)
COMMENT Contact: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cdna@resgen.com
5 prime sequence.
Location/Qualifiers
1. 170
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="cSTA20P1"
/clone_lib="potato stolon, Cornell University"
/tissue_type="axillary buds of stem explants, swelling
stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; RNA was supplied by Christian Bachem & Beatrix
Horvath(Laboratory of Plant Breeding, Dept. of Plant
Sciences, Wageningen University, The Netherlands). Total
RNA was isolated from developing axillary buds of potato
nodal stem cuttings cultured on medium for the
introduction of tuber formation as described in Bachem et
al. (Plant Journal 1996). Tissue samples were taken of
stages corresponding to growing stolons and the early

```

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadamsetigr.org

Clones are available from Research Genetics (info@resgen.com). BAC end search page:

http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html.

Seq primer: M13-21

Class: BAC ends.

#### FEATURES

source

Location/Qualifiers

1..452

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="2383N20"

/clone\_lib="CIT-HSP"

/sex="Male"

/cell\_type="Sperm"

/note="vector: pBelOBAC11; Site\_1: HindIII; Site\_2: HindIII"

BASE COUNT 128 a 88 c 88 g 148 t

#### ORIGIN

Query Match 7.8%; Score 95; DB 225; Length 452;  
Best Local Similarity 100.0%; Pred. No. 1.4e-40;  
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 588 agattacagagcacttcagagcaacatgaggtttgttaagaacaacaacacccattg 647

Db 189 AGATTACAGAAGCACTTCAGAGCAACATGAGGTTTGTGAAGAACACAAACACCCATTG 248

Qy 648 agctgtctctccaggcactgtcatgccacagcg 682

Db 249 ACCTGTCTCTCCAGGCACTGTCTATGCCACGCG 283

#### RESULT 5

AQ267168/c

LOCUS

DEFINITION RPC11-73A7.TJ RPC1-11 Homo sapiens genomic clone RPC1-11-73A7, DNA

sequence.

ACCESSION AQ267168

VERSION AQ267168.1

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1998)

Other\_GSSs: RPC11-73A7.TK

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadamsetigr.org

Clones are derived from the human BAC library RPC1-11. For BAC

library availability, please contact Pieter de Jong

(pieterdejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

Research Genetics (info@resgen.com). BAC end search page:

http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1..438

/organism="Homo sapiens"

/db\_xref="GDB:7527654"

/db\_xref="taxon:9606"

/clone="RPC1-11-73A7"

/clone\_lib="RPC1-11"

/sex="Male"

/cell\_type="Lymphocytes"

/note="vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI; RPC111 Human Male BAC Library"

BASE COUNT 153 a 81 c 78 g 126 t

#### ORIGIN

Query Match 4.0%; Score 48; DB 226; Length 438;  
Best Local Similarity 100.0%; Pred. No. 9e-15;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 799 agttccatcaagctccgcctcagagaatgctattggtgatccataaa 846

Db 48 AGTTCCATCAAGCTCCGCTCAGAGAATGCTATTGGTGATCCATATAA 1

#### RESULT 6

B92720

LOCUS

DEFINITION CIT-HSP-2170M15.TF CIT-HSP Homo sapiens genomic clone 2170M15, DNA

sequence.

ACCESSION B92720

VERSION B92720.1

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1998)

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadamsetigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html

Seq primer: M13-21;

Class: BAC ends.

Location/Qualifiers

1..354

/organism="Homo sapiens"

/db\_xref="GDB:7103209"

/db\_xref="taxon:9606"

/clone="2170M15"

/clone\_lib="CIT-HSP"

/sex="Male"

/cell\_type="Sperm"

/note="vector: pBelOBAC11; Site\_1: HindIII; Site\_2: HindIII"

BASE COUNT 93 a 70 c 67 g 124 t

#### ORIGIN

Query Match 3.8%; Score 46; DB 257; Length 354;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 588 agattacagagcacttcagagcaacatgaggtttgttaagaaca 633

Db 200 AGATTACAGAAGCACTTCAGAGCAACATGAGGTTTGTGTAAGAACA 245

```

VERSION      AA417643.1  GI:2079462
KEYWORDS     EST.
SOURCE       Homo sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 490)
AUTHORS      Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
              Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
              , J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
              White, Y., Wylie, T., Waterston, R. and Wilson, R.
              WashU-NCI human EST Project
TITLE        Unpublished (1997)
JOURNAL      Contact: Wilson RK
COMMENT       Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              This clone is available royalty-free through LLNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              Insert Length: 950 Std Error: 0.00
              Seq primer: -28ml3 rev2 ET from Amersham
              High quality sequence stop: 465.
FEATURES     Location/Qualifiers
              source
                1..490
                /organism="Homo sapiens"
                /db_xref="GDB:5975465"
                /db_xref="taxon:9606"
                /clone="IMAGE:752630"
                /clone_lib="Soares_NHMPu_S1"
                /tissue_type="Pooled human melanocyte, fetal heart, and
                pregnant uterus"
                /lab_host="DH10B"
                /note="Organ: mixed (see below); Vector: pMT3D-Pac
                (Pharmacia) with a modified polylinker; Site:1: Not I;
                Site_2: Eco RI; Equal amounts of plasmid DNA from three
                normalized libraries (melanocyte 2NbHM, pregnant uterus
                NbHPU, and fetal heart NbHH19) were mixed, and ss circles
                were made in vitro. Following HAP purification, this DNA
                was used as tracer in a subtractive hybridization
                reaction. The driver was PCR-amplified cDNAs from pools of
                5,000 clones made from the same 3 libraries. The pools
                consisted of I.M.A.G.E. clones 260232-265223,
                340488-345479, and 484488-489479."
BASE COUNT   158 a 114 c 97 g 121 t
ORIGIN
Query Match 15.6%; Score 189; DB 7; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.8e-92;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1077 cagcgtgctacacatcccaactgtctgcacagtttggcgctattaaagcacatg 1086
      |||||||
Db 35 CAGAGTGCATACATACATCCCACTGTCTGCACAGTTTGGCGGTATTAAGACGCAAT 94
      |||||||
QY 1087 ttgaaattcacactctcaaccactactggcgcgagtttgacaaatctaggaaagtggg 1146
      |||||||
Db 95 TTTGAATTACACTCTCAACCACTACTGGCGCGAGTTTGACAACTAGGAAAGTGGG 154
      |||||||
QY 1147 acaagaatgatttgaagctccaccacaaagacctaataatcaaaagattggcattg 1206
      |||||||
Db 155 ACAAGAATGATTTTGAAGACTCCACCAAGACCTAATATCAAAAGAGTTGGCATGTT 214
      |||||||
QY 1207 tggcttctg 1215
      |||||||
Db 215 TGGCTTCTG 223

RESULT      3
AW961069
LOCUS
Query Match 15.6%; Score 189; DB 7; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.8e-92;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1077 cagcgtgctacacatcccaactgtctgcacagtttggcgctattaaagcacatg 1086
      |||||||
Db 35 CAGAGTGCATACATACATCCCACTGTCTGCACAGTTTGGCGGTATTAAGACGCAAT 94
      |||||||
QY 1087 ttgaaattcacactctcaaccactactggcgcgagtttgacaaatctaggaaagtggg 1146
      |||||||
Db 95 TTTGAATTACACTCTCAACCACTACTGGCGCGAGTTTGACAACTAGGAAAGTGGG 154
      |||||||
QY 1147 acaagaatgatttgaagctccaccacaaagacctaataatcaaaagattggcattg 1206
      |||||||
Db 155 ACAAGAATGATTTTGAAGACTCCACCAAGACCTAATATCAAAAGAGTTGGCATGTT 214
      |||||||
QY 1207 tggcttctg 1215
      |||||||
Db 215 TGGCTTCTG 223

DEFINITION   EST373036 MAGE resequences, MAGF Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW961069
VERSION     1
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE    1 (bases 1 to 617)
AUTHORS      Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt
              , I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
              Quackenbush, J.
TITLE        Assessment of gene expression patterns in a model of colon tumor
              metastasis using a 19,200 element cDNA microarray
JOURNAL      Unpublished (2000)
COMMENT      Contact: John Quackenbush
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 3528
              Fax: 301 838 0208
              Email: johnq@tigr.org
              Plate: 152
              Seq primer: Forward.
              Location/Qualifiers
                source
                  1..617
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone_lib="MAGE resequences, MAGF"
                  /note="Vector: pBluescriptSKm"
BASE COUNT   211 a 116 c 127 g 163 t
ORIGIN
Query Match 10.4%; Score 126; DB 122; Length 617;
Best Local Similarity 99.4%; Pred. No. 1.4e-57;
Matches 176; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 9gtctcttgagtgctcttctcactaagtcccgagagctggaagccatcaagttaatt 60
      |||||||
Db 441 GGTCTCTTGAAGTGCTCTTCTTCACTAAGTGCCGAGGAGCTGGAAGGCATCAAGTTAAT 500
      |||||||
QY 61 tagatttggcaaaaagaaaaatacacattcaactgttgtttttataatccctgaaaaattt 120
      |||||||
Db 501 TAGATTGGCAAAAAGAAAAATACACATTCTCTGTTGTTTATAATCCCTGAAAAATTT 560
      |||||||
QY 121 taaaggttgatttcaggcgcatggaatggatattgcttcaactgaaccactgacaat 177
      |||||||
Db 561 TAAAGTTGTATTTTCAGGGCATGGGATGGATATTTGCTTTAACTGAACCACTGACAAAT 617
      |||||||

RESULT      4
AW961069
LOCUS
DEFINITION   CIT-HSP-2383N20.TF CIT-HSP Homo sapiens genomic clone 2383N20, DNA
              sequence.
ACCESSION   AQ196491
VERSION     AQ196491
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE    1 (bases 1 to 452)
AUTHORS      Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
              Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and
              Venter, J.C.
TITLE        Use of a random human BAC End Sequence Database for Sequence-Ready
              Map Building
JOURNAL      Unpublished (1998)
COMMENT      Other GSSs: CIT-HSP-2383N20.TR
              Contact: Mark Adams
              Department of Eukaryotic Genomics
              The Institute for Genomic Research

```





117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: em\_est100:\*  
170: em\_est101:\*  
171: em\_est102:\*  
172: em\_est103:\*  
173: em\_est104:\*  
174: em\_est105:\*  
175: em\_est106:\*  
176: em\_est107:\*  
177: em\_est108:\*  
178: em\_est109:\*  
179: em\_est110:\*  
180: em\_est111:\*  
181: em\_est112:\*  
182: em\_est113:\*  
183: em\_est114:\*  
184: em\_est115:\*  
185: em\_est116:\*  
186: em\_est117:\*  
187: em\_est118:\*  
188: em\_est119:\*  
189: em\_est120:\*

190: gb\_est121:\*  
191: gb\_est122:\*  
192: gb\_est123:\*  
193: em\_gss\_fun1:\*  
194: em\_gss\_fun2:\*  
195: em\_gss\_fun3:\*  
196: em\_gss\_fun4:\*  
197: em\_gss\_fun5:\*  
198: em\_gss\_fun6:\*  
199: em\_gss\_fun7:\*  
200: em\_gss\_fun8:\*  
201: em\_gss\_fun9:\*  
202: em\_gss\_inv1:\*  
203: em\_gss\_inv2:\*  
204: em\_gss\_inv3:\*  
205: em\_gss\_inv4:\*  
206: em\_gss\_inv5:\*  
207: em\_gss\_inv6:\*  
208: em\_gss\_inv7:\*  
209: em\_gss\_inv8:\*  
210: em\_gss\_inv9:\*  
211: em\_gss\_inv10:\*  
212: em\_gss\_inv11:\*  
213: em\_gss\_inv12:\*  
214: em\_gss\_inv13:\*  
215: em\_gss\_inv14:\*  
216: em\_gss\_inv15:\*  
217: em\_gss\_inv16:\*  
218: em\_gss\_inv17:\*  
219: em\_gss\_inv18:\*  
220: em\_gss\_inv19:\*  
221: em\_gss\_inv20:\*  
222: em\_gss\_inv21:\*  
223: em\_gss\_inv22:\*  
224: em\_gss\_inv23:\*  
225: em\_gss\_inv24:\*  
226: em\_gss\_inv25:\*  
227: em\_gss\_inv26:\*  
228: em\_gss\_inv27:\*  
229: em\_gss\_inv28:\*  
230: em\_gss\_inv29:\*  
231: em\_gss\_inv30:\*  
232: em\_gss\_inv31:\*  
233: em\_gss\_inv32:\*  
234: em\_gss\_inv33:\*  
235: em\_gss\_inv34:\*  
236: em\_gss\_inv35:\*  
237: em\_gss\_inv36:\*  
238: em\_gss\_inv37:\*  
239: em\_gss\_inv38:\*  
240: em\_gss\_inv39:\*  
241: em\_gss\_inv40:\*  
242: em\_gss\_inv41:\*  
243: em\_gss\_inv42:\*  
244: em\_gss\_inv43:\*  
245: em\_gss\_inv44:\*  
246: em\_gss\_inv45:\*  
247: em\_gss\_inv46:\*  
248: em\_gss\_inv47:\*  
249: em\_gss\_inv48:\*  
250: em\_gss\_inv49:\*  
251: em\_gss\_inv50:\*  
252: em\_gss\_inv51:\*  
253: em\_gss\_inv52:\*  
254: em\_gss\_inv53:\*  
255: em\_gss\_inv54:\*  
256: em\_gss\_inv55:\*  
257: em\_gss\_inv56:\*  
258: em\_gss\_inv57:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2001, 10:56:30 ; Search time 1299.33 Seconds  
(without alignments)  
8839.338 Million cell updates/sec

Title: US-09-389-000-1\_COPY\_735\_1949

Perfect score: 1215

Sequence: 1 ggtctcttgaagggtgcctt.....ttggcatggttggtctctg 1215

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 10228115 seqs, 4726426750 residues

Word size : 0

Total number of hits satisfying chosen parameters: 20383376

Minimum DB seq length: 0

Maximum DB seq length: 1214

Post-processing: Listing first 45 summaries

Database :

EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_est3:\*
- 4: gb\_est4:\*
- 5: gb\_est5:\*
- 6: gb\_est6:\*
- 7: gb\_est7:\*
- 8: gb\_est8:\*
- 9: gb\_est9:\*
- 10: gb\_est10:\*
- 11: gb\_est11:\*
- 12: gb\_est12:\*
- 13: gb\_est13:\*
- 14: gb\_est14:\*
- 15: gb\_est15:\*
- 16: gb\_est16:\*
- 17: gb\_est17:\*
- 18: gb\_est18:\*
- 19: gb\_est19:\*
- 20: gb\_est20:\*
- 21: gb\_est21:\*
- 22: gb\_est22:\*
- 23: gb\_est23:\*
- 24: gb\_est24:\*
- 25: gb\_est33:\*
- 26: gb\_est34:\*
- 27: gb\_est35:\*
- 28: gb\_est36:\*
- 29: gb\_est37:\*
- 30: gb\_est38:\*
- 31: gb\_est39:\*
- 32: gb\_est40:\*
- 33: em\_estba:\*
- 34: em\_estfun:\*
- 35: em\_esthum1:\*
- 36: em\_esthum2:\*
- 37: em\_esthum3:\*
- 38: em\_esthum4:\*
- 39: em\_esthum5:\*
- 40: em\_esthum6:\*
- 41: em\_esthum7:\*
- 42: em\_esthum8:\*
- 43: em\_esthum9:\*

- 44: em\_esthum10:\*
- 45: em\_esthum11:\*
- 46: em\_esthum12:\*
- 47: em\_esthum13:\*
- 48: em\_esthum14:\*
- 49: em\_esthum15:\*
- 50: em\_esthum16:\*
- 51: em\_esthum17:\*
- 52: em\_esthum18:\*
- 53: em\_esthum19:\*
- 54: em\_esthum20:\*
- 55: em\_esthum21:\*
- 56: em\_esthum22:\*
- 57: em\_esthum23:\*
- 58: em\_esthum24:\*
- 59: em\_esthum25:\*
- 60: em\_esthum26:\*
- 61: em\_esthum27:\*
- 62: em\_esthum28:\*
- 63: em\_estin1:\*
- 64: em\_estin2:\*
- 65: em\_estin3:\*
- 66: em\_estin4:\*
- 67: em\_estin5:\*
- 68: em\_estom1:\*
- 69: em\_estom2:\*
- 70: em\_estov1:\*
- 71: em\_estov2:\*
- 72: em\_estpl1:\*
- 73: em\_estpl2:\*
- 74: em\_estpl3:\*
- 75: em\_estpl4:\*
- 76: em\_estpl5:\*
- 77: em\_estpl6:\*
- 78: em\_estpl7:\*
- 79: em\_estpl8:\*
- 80: em\_estpl9:\*
- 81: em\_estpl10:\*
- 82: em\_estro1:\*
- 83: em\_estro2:\*
- 84: em\_estro3:\*
- 85: em\_estro4:\*
- 86: em\_estro5:\*
- 87: em\_estro6:\*
- 88: em\_estro7:\*
- 89: em\_estro8:\*
- 90: em\_estro9:\*
- 91: em\_estro10:\*
- 92: em\_estro11:\*
- 93: em\_estro12:\*
- 94: em\_estro13:\*
- 95: em\_estro14:\*
- 96: em\_estro15:\*
- 97: em\_estro16:\*
- 98: em\_estro17:\*
- 99: em\_estro18:\*
- 100: em\_estro19:\*
- 101: em\_estro20:\*
- 102: gb\_est25:\*
- 103: gb\_est26:\*
- 104: gb\_est27:\*
- 105: gb\_est28:\*
- 106: gb\_est29:\*
- 107: gb\_est30:\*
- 108: gb\_est31:\*
- 109: gb\_est32:\*
- 110: gb\_est41:\*
- 111: gb\_est42:\*
- 112: gb\_est43:\*
- 113: gb\_est44:\*
- 114: gb\_est45:\*
- 115: gb\_est46:\*
- 116: gb\_est47:\*

**THIS PAGE BLANK (USPTO)**

Qy 949 gaacatttcaattca 963  
          |||||  
Db 16 GAACATTTCaATTCA 2

Search completed: September 15, 2001, 12:35:52  
Job time: 3697 sec

Query Match 1.2%; Score 15; DB 1; Length 31;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 949 gaacatttcattca 963  
DB 16 GAACATTTCAATTCA 2

RESULT 13  
US-08-435-634-306/c  
; Sequence 306, Application US/08435634  
; Patent No. 5731295  
; GENERAL INFORMATION:  
; APPLICANT: Draper, Kenneth G.  
; APPLICANT: Pavco, Pamela  
; APPLICANT: McSwiggen, James  
; APPLICANT: Gustofson, John  
; APPLICANT: Stinchcomb, Dan T.  
; TITLE OF INVENTION: METHOD AND REAGENT FOR TREATMENT  
; OF ARTHRITIC CONDITIONS  
; NUMBER OF SEQUENCES: 1151  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/435,634  
; FILING DATE: 05-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/390,850  
; FILING DATE: February 17, 1995  
; APPLICATION NUMBER: 08/354,920  
; FILING DATE: December 13, 1994  
; APPLICATION NUMBER: 08/152,487  
; FILING DATE: No. 5731295ember 12, 1993  
; APPLICATION NUMBER: 07/989,848  
; FILING DATE: December 7, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 211/084  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 306:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-435-634-306

Query Match 1.2%; Score 15; DB 1; Length 31;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 949 gaacatttcattca 963  
DB 25 GAACATTTCAATTCA 11

RESULT 14  
US-08-435-634-307/c  
; Sequence 307, Application US/08435634  
; Patent No. 5731295  
; GENERAL INFORMATION:  
; APPLICANT: Draper, Kenneth G.  
; APPLICANT: Pavco, Pamela  
; APPLICANT: McSwiggen, James  
; APPLICANT: Gustofson, John  
; APPLICANT: Stinchcomb, Dan T.  
; TITLE OF INVENTION: METHOD AND REAGENT FOR TREATMENT  
; OF ARTHRITIC CONDITIONS  
; NUMBER OF SEQUENCES: 1151  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/435,634  
; FILING DATE: 05-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/390,850  
; FILING DATE: February 17, 1995  
; APPLICATION NUMBER: 08/354,920  
; FILING DATE: December 13, 1994  
; APPLICATION NUMBER: 08/152,487  
; FILING DATE: No. 5731295ember 12, 1993  
; APPLICATION NUMBER: 07/989,848  
; FILING DATE: December 7, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 211/084  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 307:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-435-634-307

Query Match 1.2%; Score 15; DB 1; Length 31;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 949 gaacatttcattca 963  
DB 17 GAACATTTCAATTCA 3

RESULT 15  
US-08-435-634-308/c  
; Sequence 308, Application US/08435634  
; Patent No. 5731295  
; GENERAL INFORMATION:

; FILING DATE: No. 5612215ember 12, 1993  
; APPLICATION NUMBER: 07/989,848  
; FILING DATE: December 7, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 211/084  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 306:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-390-850-306

Query Match 1.2%; Score 15; DB 1; Length 31;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0;

Qy 949 gaacattcaattca 963  
|||||  
Db 25 GAACATTTCAATTCA 11

RESULT 11  
US-08-390-850-307/c  
; Sequence 307, Application US/08390850  
; Patent No. 5612215  
; GENERAL INFORMATION:  
; APPLICANT: Draper, Kenneth G.  
; APPLICANT: Pavco, Pamela  
; APPLICANT: McSwiggen, James  
; APPLICANT: Gustofson, John  
; APPLICANT: Stinchcomb, Dan T.  
; TITLE OF INVENTION: METHOD AND REAGENT FOR TREATMENT  
; OF ARTHRITIC CONDITIONS  
; NUMBER OF SEQUENCES: 1151  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/390,850  
; FILING DATE: February 17, 1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/354,920  
; FILING DATE: December 13, 1994  
; APPLICATION NUMBER: 08/152,487  
; FILING DATE: December 7, 1992  
; APPLICATION NUMBER: No. 5612215ember 12, 1993  
; FILING DATE: December 7, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 211/084  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440

; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 307:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-390-850-307

Query Match 1.2%; Score 15; DB 1; Length 31;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0;

Qy 949 gaacattcaattca 963  
|||||  
Db 17 GAACATTTCAATTCA 3

RESULT 12  
US-08-390-850-308/c  
; Sequence 308, Application US/08390850  
; Patent No. 5612215  
; GENERAL INFORMATION:  
; APPLICANT: Draper, Kenneth G.  
; APPLICANT: Pavco, Pamela  
; APPLICANT: McSwiggen, James  
; APPLICANT: Gustofson, John  
; APPLICANT: Stinchcomb, Dan T.  
; TITLE OF INVENTION: METHOD AND REAGENT FOR TREATMENT  
; OF ARTHRITIC CONDITIONS  
; NUMBER OF SEQUENCES: 1151  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/390,850  
; FILING DATE: February 17, 1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/354,920  
; FILING DATE: December 13, 1994  
; APPLICATION NUMBER: 08/152,487  
; FILING DATE: December 7, 1992  
; APPLICATION NUMBER: No. 5612215ember 12, 1993  
; FILING DATE: December 7, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 211/084  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 308:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-390-850-308

APPLICANT: Li, Linheng  
APPLICANT: Hood, Leroy  
APPLICANT: Krantz, Ian D.  
APPLICANT: Spinner, Nancy B.  
TITLE OF INVENTION: Human Jagged Polypeptide, Encoding  
TITLE OF INVENTION: Nucleic Acids and Methods of Use  
NUMBER OF SEQUENCES: 110  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/882,046  
FILING DATE: 25-JUN-1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-UW 2637  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: intron  
LOCATION: 1..10  
NAME/KEY: exon  
LOCATION: 11..20  
US-08-882-046-28

Query Match 1.28; Score 15; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 276 tgcagaggctctaca 290  
|||||  
Db 6 TGCAGAGGCTCTACA 20

RESULT 9  
US-08-859-998-622  
Sequence 622, Application US/0885998  
Patent No. 5994076  
GENERAL INFORMATION:  
APPLICANT: Chenchik, Alex  
APPLICANT: Jokhadze, George  
APPLICANT: Biblasiavili, Robert  
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 1375  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 2200 Sand Hill Road, Suite 100  
CITY: Menlo Park  
STATE: CA  
COUNTRY: US  
ZIP: 94025  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/859,998  
FILING DATE: 21-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Field, Bret E.  
REGISTRATION NUMBER: 37,620  
REFERENCE/DOCKET NUMBER: 09096/002001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-322-5070  
TELEFAX: 415-854-0875  
INFORMATION FOR SEQ ID NO: 622:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
OTHER INFORMATION: oligonucleotide primer  
US-08-859-998-622  
Query Match 1.28; Score 15; DB 2; Length 28;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 657 tccaggcactgtca 671  
|||||  
Db 11 TCCCAGGCACTGTCA 25  
RESULT 10  
US-08-390-850-306/c  
Sequence 306, Application US/08390850  
Patent No. 5612215  
GENERAL INFORMATION:  
APPLICANT: Draper, Kenneth G.  
APPLICANT: Pavco, Pamela  
APPLICANT: McSwiggen, James  
APPLICANT: Gustofson, John  
APPLICANT: Stinchcomb, Dan T.  
TITLE OF INVENTION: METHOD AND REAGENT FOR TREATMENT  
TITLE OF INVENTION: OF ARTHRITIC CONDITIONS  
NUMBER OF SEQUENCES: 1151  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/390,850  
FILING DATE: February 17, 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/354,920  
FILING DATE: December 13, 1994  
APPLICATION NUMBER: 08/152,487

APPLICANT: DUKTA-MALEN, Sylvie  
APPLICANT: CHARLES, Murielle  
APPLICANT: EVERS, Stefan  
APPLICANT: CASADEWALL, Barbara  
TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR  
TITLE OF INVENTION: DETECTING ENTEROCOCCI AND STREPTOCOCCI BACTERIAL STRAINS  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner, L.L.P.  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/882,501  
FILING DATE: 25-JUN-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McDonnell, Leslie A.  
REGISTRATION NUMBER: 34,872  
REFERENCE/DOCKET NUMBER: 03495.0155-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1095 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus hirae  
US-08-882-501-31

Query Match 1.3%; Score 16; DB 3; Length 1095;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 231 ttaagactgaaacgc 246  
|||||  
Db 90 TTAAGACTGAAACGC 75

RESULT 6  
US-08-451-691-3  
Sequence 3, Application US/08451691  
Patent No. 5557032  
GENERAL INFORMATION:  
APPLICANT: Tak W. Mak et al.  
TITLE OF INVENTION: Knockout Mammals  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: Amgen Center  
STREET: 1840 Dehavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in., DS, 2.0 MB  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh OS 7.0.  
SOFTWARE: Microsoft Word Version 5.1a

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/451,691  
FILING DATE:  
CLASSIFICATION: 800  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 Base pairs  
TYPE: Nucleic acid  
STRANDEDNESS: Single stranded  
TOPOLOGY: Linear  
US-08-451-691-3

Query Match 1.2%; Score 15; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1036 actacacatcccaac 1050  
|||||  
Db 4 ACTACATCCCAAC 18

RESULT 7  
US-08-528-363-3  
Sequence 3, Application US/08528363  
Patent No. 5616491  
GENERAL INFORMATION:  
APPLICANT: Tak W. Mak  
TITLE OF INVENTION: Knockout Mammals  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: Amgen Center  
STREET: 1840 Dehavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in., DS, 2.0 MB  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh OS 7.0.  
SOFTWARE: Microsoft Word Version 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/528,363  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/067,767  
FILING DATE: 26-MAY-1993  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 Base pairs  
TYPE: Nucleic acid  
STRANDEDNESS: Single stranded  
TOPOLOGY: Linear  
US-08-528-363-3

Query Match 1.2%; Score 15; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1036 actacacatcccaac 1050  
|||||  
Db 4 ACTACATCCCAAC 18

RESULT 8  
US-08-882-046-28  
Sequence 28, Application US/08882046  
Patent No. 6136952  
GENERAL INFORMATION:



```
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896.164
; FILING DATE: July 17, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6218521man D. Hanson
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-896-164-47

Query Match 1.3%; Score 16; DB 4; Length 822;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 255 ctgaagaacttgatt 270
Db 592 CTGAAGAAGCTGGATT 577

RESULT 3
US-08-162-475A-3/c
; Sequence 3, Application US/08162475A
; Patent No. 5656474
; GENERAL INFORMATION:
; APPLICANT: Zohreh Tabaeizadeh
; TITLE OF INVENTION: A novel endochitinase gene
; TITLE OF INVENTION: Induced by osmotic stress and abscisic acid isolated
; TITLE OF INVENTION: from the wild tomato Lycopersicon chilense Dun.
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 New York Avenue, N.W., 9th Floor
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 720 kb diskette
; COMPUTER: IBM PS/2, Model 30
; OPERATING SYSTEM: PC-DOS 3.30
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,475A
; FILING DATE: December 7, 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, Paul E. Jr.
; REGISTRATION NUMBER: 32011
; REFERENCE/DOCKET NUMBER: PEW/3122/204351
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE: Lycopersicon chilense
;
US-08-162-475A-3

Query Match 1.3%; Score 16; DB 4; Length 966;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 314 actgatctattgcct 329
Db 360 ACTGATCTATTGCCT 345

RESULT 5
US-08-882-501-31/c
; Sequence 31, Application US/08882501
; Patent No. 6054269
; GENERAL INFORMATION:
; APPLICANT: GARNIER, Fabien
; APPLICANT: GERBAUD, Guy
; APPLICANT: GALIMAND, Marc
; APPLICANT: COURVALIN, Patrice
```

```
Query Match 1.3%; Score 16; DB 1; Length 935;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 314 actgatctattgcct 329
Db 331 ACTGATCTATTGCCT 316

RESULT 4
US-08-162-475A-1/c
; Sequence 1, Application US/08162475A
; Patent No. 5656474
; GENERAL INFORMATION:
; APPLICANT: Zohreh Tabaeizadeh
; TITLE OF INVENTION: A novel endochitinase gene
; TITLE OF INVENTION: Induced by osmotic stress and abscisic acid isolated
; TITLE OF INVENTION: from the wild tomato Lycopersicon chilense Dun.
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 New York Avenue, N.W., 9th Floor
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 720 kb diskette
; COMPUTER: IBM PS/2, Model 30
; OPERATING SYSTEM: PC-DOS 3.30
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,475A
; FILING DATE: December 7, 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, Paul E. Jr.
; REGISTRATION NUMBER: 32011
; REFERENCE/DOCKET NUMBER: PEW/3122/204351
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE: Lycopersicon chilense
;
US-08-162-475A-1

Query Match 1.3%; Score 16; DB 1; Length 966;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 314 actgatctattgcct 329
Db 360 ACTGATCTATTGCCT 345

RESULT 5
US-08-882-501-31/c
; Sequence 31, Application US/08882501
; Patent No. 6054269
; GENERAL INFORMATION:
; APPLICANT: GARNIER, Fabien
; APPLICANT: GERBAUD, Guy
; APPLICANT: GALIMAND, Marc
; APPLICANT: COURVALIN, Patrice
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2001, 11:34:15 : Search time 79.34 Seconds  
(without alignments)  
2899.080 Million cell updates/sec

Title: US-09-389-000-1\_COPY\_735\_1949

Perfect score: 1215

Sequence: 1 ggtctcttgaaggcctt.....ttggcatggttggtcttg 1215

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 324599 seqs, 94655562 residues

Word size : 0

Total number of hits satisfying chosen parameters: 605014

Minimum DB seq length: 0

Maximum DB seq length: 1214

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgnl\_7/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgnl\_7/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgnl\_7/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgnl\_7/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgnl\_7/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgnl\_7/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	16	1.3	422	3	US-08-826-532-15
C 2	16	1.3	822	4	US-08-896-164-47
C 3	16	1.3	935	1	US-08-162-475A-3
C 4	16	1.3	966	1	US-08-162-475A-1
C 5	16	1.3	1095	3	US-08-882-501-31
C 6	15	1.2	20	1	US-08-451-691-3
C 7	15	1.2	20	1	US-08-528-363-3
C 8	15	1.2	20	3	US-08-882-046-28
C 9	15	1.2	28	2	US-08-859-998-622
C 10	15	1.2	31	1	US-08-330-850-306
C 11	15	1.2	31	1	US-08-330-850-307
C 12	15	1.2	31	1	US-08-330-850-308
C 13	15	1.2	31	1	US-08-435-634-306
C 14	15	1.2	31	1	US-08-435-634-307
C 15	15	1.2	31	1	US-08-435-634-308
C 16	15	1.2	83	6	5240846-1
C 17	15	1.2	247	1	US-08-691-641-4
C 18	15	1.2	372	1	US-08-630-822A-96
C 19	15	1.2	372	2	US-09-005-069-96
C 20	15	1.2	380	1	US-08-270-076A-3
C 21	15	1.2	516	4	US-09-276-531-124
C 22	15	1.2	522	6	5240847-5
C 23	15	1.2	535	6	5240847-6
C 24	15	1.2	594	4	US-09-075-019-1
C 25	15	1.2	600	6	5240847-8
C 26	15	1.2	600	6	5240847-9
C 27	15	1.2	600	6	5240847-16

Patent No. 5240847  
Sequence 7, Appli  
Sequence 8, Appli  
Sequence 752, App  
Patent No. 5240847  
Sequence 29, Appli  
Sequence 6, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 13, Appli  
Sequence 13, Appli  
Sequence 15, App  
Sequence 1, Appli  
Sequence 1, Appli

28 15 1.2 600 6 5240847-17  
29 15 1.2 609 2 US-08-308-952-7  
30 15 1.2 609 4 US-08-124-141-8  
31 15 1.2 612 4 US-08-998-416-752  
32 15 1.2 623 6 5240847-1  
33 15 1.2 678 3 US-09-181-183-29  
34 15 1.2 681 4 US-09-075-019-6  
35 15 1.2 719 1 US-08-375-346A-1  
36 15 1.2 719 2 US-08-467-123B-1  
37 15 1.2 724 2 US-08-825-780-2  
38 15 1.2 739 1 US-08-156-383-3  
39 15 1.2 739 5 PCT-US92-10866-3  
40 15 1.2 739 5 PCT-US92-10866-3  
41 15 1.2 763 2 US-08-544-822-13  
42 15 1.2 763 3 US-09-070-964-13  
43 15 1.2 784 3 US-08-961-083-151  
44 15 1.2 813 2 US-08-927-230A-1  
45 15 1.2 813 3 US-09-151-052-1

#### ALIGNMENTS

RESULT 1  
US-08-826-532-15/c  
; Sequence 15, Application US/08826532B  
; Patent No. 6027923  
; GENERAL INFORMATION:  
; APPLICANT: Wallace, Robert B.  
; TITLE OF INVENTION: Linked Linear Amplification of Nucleic Acids  
; FILE REFERENCE: 3239-102P  
; CURRENT APPLICATION NUMBER: US/08/826,532B  
; CURRENT FILING DATE: 1997-04-02  
; EARLIER APPLICATION NUMBER: US 08/475,605  
; EARLIER FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 422  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-08-826-532-15

Query Match 1.3%; Score 16; DB 3; Length 422;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 908 aaagtcaccccttccct 923  
|||||  
Db 146 AAAGTCACCCCTTCCT 131

RESULT 2  
US-08-896-164-47/c  
; Sequence 47, Application US/08896164  
; Patent No. 6218521  
; GENERAL INFORMATION:  
; APPLICANT: OBATA, Yuichi  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED  
; TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR  
; TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER  
; NUMBER OF SEQUENCES: 87  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb  
; COMPUTER: IBM PS/2

**THIS PAGE BLANK (USPTO)**



PS Claim 1; Page 2844; 3271pp; English.  
XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
CC of the invention. The DNA sequences are recorded on a computer readable  
CC medium, preferably selected from a floppy or hard disk, random access  
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
CC the S.aureus DNA sequences allows putative functions to be assigned so  
CC that protein-encoding or regulatory regions of commercial, therapeutic or  
CC industrial importance can be obtained. Specifically, sequences which are  
CC likely to encode antigens have been identified and these polypeptides can  
CC be used in a vaccine composition against S.aureus infection. The  
CC polypeptides can also be used in a kit for the immunodetection of  
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,  
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
CC syndrome, etc. Organisms transformed with the DNA sequences can be used  
CC for recombinant production of the polypeptides. The new DNA sequences  
CC (and their fragments) are useful as primers or probes for isolating  
CC homologues of any of the S.aureus DNA sequences contained on the  
CC computer readable medium.  
XX  
SQ Sequence 370 BP; 118 A; 67 C; 87 G; 95 T; 3 other;

Query Match 1.4%; Score 17; DB 18; Length 370;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 153 ttgcttaactgaacca 169  
|||||  
DB 68 ttgcttaactgaacca 84

RESULT 15  
AAV78096/C  
ID AAV78096 standard; DNA: 400 BP.  
XX  
AC AAV78096;  
XX  
DT 16-MAR-1999 (first entry)  
XX  
DE Staphylococcus aureus contig SEQ ID #3785.  
XX  
KM Computer readable medium: vaccine; S.aureus infection; immunodetection;  
KM cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;  
KM skin infection; surgical wound infection; scalded skin syndrome;  
KM toxic shock syndrome; ds.  
XX  
OS Staphylococcus aureus.  
XX  
PN EP786519-A2.  
XX  
PD 30-JUL-1997.  
XX  
PF 07-JAN-1997; 97EP-0100117.  
XX  
PR 05-JAN-1996; 96US-0009861.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;  
PI Rosen CA;  
XX  
DR WPI: 1997-374922/35.  
XX  
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus  
PT stored on computer readable medium and used in the production of  
PT anti-S.aureus vaccines  
XX  
PS Claim 1; Page 2687; 3271pp; English.  
XX  
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
CC of the invention. The DNA sequences are recorded on a computer readable

CC medium, preferably selected from a floppy or hard disk, random access  
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
CC the S.aureus DNA sequences allows putative functions to be assigned so  
CC that protein-encoding or regulatory regions of commercial, therapeutic or  
CC industrial importance can be obtained. Specifically, sequences which are  
CC likely to encode antigens have been identified and these polypeptides can  
CC be used in a vaccine composition against S.aureus infection. The  
CC polypeptides can also be used in a kit for the immunodetection of  
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,  
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
CC syndrome, etc. Organisms transformed with the DNA sequences can be used  
CC for recombinant production of the polypeptides. The new DNA sequences  
CC (and their fragments) are useful as primers or probes for isolating  
CC homologues of any of the S.aureus DNA sequences contained on the  
CC computer readable medium.  
XX  
SQ Sequence 400 BP; 117 A; 83 C; 81 G; 118 T; 1 other;

Query Match 1.4%; Score 17; DB 18; Length 400;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 153 ttgcttaactgaacca 169  
|||||  
DB 140 TTGCTTTACTGACCA 124

Search completed: September 15, 2001, 12:38:12  
Job time: 3727 sec

```

XX EP1033401-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 21-FEB-2000; 2000EP-0200610.
PF
XX
XX 26-FEB-1999; 99US-0122487.
PR
XX
XX (GEST ) GENSET.
PA
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI: 2000-500381/45.
DR
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PR diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1: SEQ ID 32607; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
XX Sequence 341 BP; 154 A; 42 C; 61 G; 79 T; 5 other;
SQ
Query Match 1.4%; Score 17; DB 21; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 107 atccctgaaattttaa 123
Db 70 atccctgaaattttaa 86
RESULT 13
AAC21590/c
ID AAC21590 standard; cDNA; 344 BP.
XX
XX AAC21590;
AC
XX
XX 06-OCT-2000 (first entry)
DT
XX
XX Human secreted protein 5' EST, SEQ ID NO: 25665.
DE
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
OS
XX
XX EP1033401-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 21-FEB-2000; 2000EP-0200610.
PF
XX
XX 26-FEB-1999; 99US-0122487.
PR
XX
XX (GEST ) GENSET.
PA
XX
XX

```

```

PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI: 2000-500381/45.
DR
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PR diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1: SEQ ID 25665; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
XX Sequence 344 BP; 90 A; 59 C; 51 G; 144 T; 0 other;
SQ
Query Match 1.4%; Score 17; DB 21; Length 344;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 73 aaagaanaatcacatt 89
Db 213 AAAGAAAATACACATT 197
RESULT 14
AAV78456
ID AAV78456 standard; DNA; 370 BP.
XX
XX AAV78456;
AC
XX
XX 16-MAR-1999 (first entry)
DT
XX
XX Staphylococcus aureus contig SEQ ID #4145.
DE
XX
XX Computer readable medium; vaccine; S. aureus infection; immunodetection;
KM cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KM skin infection; surgical wound infection; scalded skin syndrome;
KM toxic shock syndrome; ds.
XX
XX Staphylococcus aureus.
OS
XX
XX EP786519-A2.
PN
XX
XX 30-JUL-1997.
PD
XX
XX 07-JAN-1997; 97EP-0100117.
PF
XX
XX 05-JAN-1996; 96US-0009861.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
PI Rosen CA;
XX
XX WPI: 1997-374922/35.
DR
XX
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus
PT stored on computer readable medium and used in the production of
PR anti-S. aureus vaccines
XX

```

KW toxic shock syndrome; ds.  
XX  
OS Staphylococcus aureus.  
XX  
PN EP786519-A2.  
XX  
PD 30-JUL-1997.  
XX  
PF 07-JAN-1997; 97EP-0100117.  
XX  
PR 05-JAN-1996; 96US-0009861.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;  
PI Rosen CA;  
XX  
DR WPI; 1997-374922/35.  
XX  
PT Polynucleotide(s) and proteins derived from Staphylococcus aureus  
PT stored on computer readable medium and used in the production of  
PT anti-S.aureus vaccines  
XX  
PS Claim 1; Page 2942; 3271pp; English.  
XX  
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences  
CC of the invention. The DNA sequences are recorded on a computer readable  
CC medium, preferably selected from a floppy or hard disk, random access  
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
CC the S.aureus DNA sequences allows putative functions to be assigned so  
CC that protein-encoding or regulatory regions of commercial, therapeutic or  
CC industrial importance can be obtained. Specifically, sequences which are  
CC likely to encode antigens have been identified and these polypeptides can  
CC be used in a vaccine composition against S.aureus infection. The  
CC polypeptides can also be used in a kit for the immunodetection of  
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,  
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
CC syndrome, etc. Organisms transformed with the DNA sequences can be used  
CC for recombinant production of the polypeptides. The new DNA sequences  
CC (and their fragments) are useful as primers or probes for isolating  
CC homologues of any of the S.aureus DNA sequences contained on the  
CC computer readable medium.  
XX  
SQ Sequence 304 BP; 81 A; 68 C; 62 G; 90 T; 3 other;

Query Match 1.4%; Score 17; DB 18; Length 304;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 153 ttgcttaactgaacca 169  
|||||  
DB 189 TTGCTTTACTGACCA 173

RESULT 11  
AAFI1660/C  
ID AAFI1660 standard; cDNA; 330 BP.  
XX  
AC AAFI1660;  
XX  
DT 13-MAR-2001 (first entry)  
XX  
DE Aspergillus niger EST SEQ ID NO:4183.  
XX  
KW Multiple gene expression; filamentous fungal cell; EST;  
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;  
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
KW culture condition; environmental stress; spore morphogenesis;  
KW metabolic pathway engineering; catabolic pathway engineering; ss.  
XX  
OS Aspergillus niger.

XX  
PN WO200056762-A2.  
XX  
PD 28-SEP-2000.  
XX  
PF 22-MAR-2000; 2000WO-US07781.  
XX  
PR 22-MAR-1999; 99US-0273623.  
XX  
PA (NOVO) NOVO NORDISK BIOTECH INC.  
XX  
PI (NOVO) NOVO NORDISK AS.  
XX  
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
XX  
DR WPI; 2000-594572/56.  
XX  
PT Monitoring differential expression of genes in filamentous fungal cells  
PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
PT substrate of expressed sequence tags -  
XX  
PS Claim 87; Page 1826; 3161pp; English.  
XX  
CC The present invention describes a method for monitoring differential  
CC expression of genes in a first filamentous fungal (FF) cell relative to  
CC expression of the same genes in one or more second filamentous fungal  
CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
CC are used in the methods for monitoring differential expression of genes  
CC in a first filamentous fungal (FF) cell relative to expression of the  
CC same genes in one or more second filamentous fungal cells. Monitoring  
CC the global expression of genes from FF cells allows the production  
CC potential of the microorganisms to be improved. New genes may be  
CC discovered, possible functions of unknown open reading frames can be  
CC identified and gene copy number variation and stability can be  
CC monitored. The expression of genes can be used to study how FF cells  
CC adapt to changes in culture conditions, environmental stress, spore  
CC morphogenesis, recombination, metabolic or catabolic pathway  
CC engineering. Using ESTs provides several advantages over genomic or  
CC random cDNA clones including elimination of redundancy as one spot on an  
CC array equals one gene or open reading frame, and organisation of the  
CC microarrays based on function of the gene products to facilitate  
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from  
CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus  
CC niger; AAF1854 to AAF14878 represents ESTs from Aspergillus oryzae; and  
CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are  
CC all specifically claimed in the present invention.  
XX  
SQ Sequence 330 BP; 67 A; 66 C; 73 G; 124 T; 0 other;

Query Match 1.4%; Score 17; DB 21; Length 330;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 aagaacaacaacacc 642  
|||||  
DB 49 AAGAAACAAACAAACACC 33

RESULT 12  
AAC28532  
ID AAC28532 standard; cDNA; 341 BP.  
XX  
AC AAC28532;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 32607.  
XX  
KW Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 400 aaagatctctctctca 417  
|||||  
Db 312 aaagatctctctctca 329

## RESULT 8

AA086825  
ID AA086825 standard; DNA; 28 BP.

AC AA086825;

DT 20-DEC-1995 (first entry)

DE Primer #2 for acetoaldehyde acid isomeroreductase gene promoter.

KM Promoter; B.flavum; acetoaldehyde acid isomeroreductase gene;  
KM plasmid vector; PCR; primer; polymerase chain; amplify; ss.

OS Synthetic.

PN JP07087977-A.

PD 04-APR-1995.

PF 12-AUG-1993; 93JP-0200707.

PR 12-AUG-1993; 93JP-0200707.

PA (MTP) MITSUBISHI PETROCHEMICAL CO LTD.

DR WPI; 1995-166380/22.

PT A DNA fragment having promoter function in Coryneform bacteria  
PT may be used for the high level expression of structural genes

PS Disclosure; Page 7; 7pp; Japanese.

CC The sequences given in AA086824-25 are primers which were used in the  
CC amplification of the promoter fragment derived from the B.flavum  
CC acetoaldehyde acid isomeroreductase gene. The amplified sequence may  
CC be used in a plasmid vector, operatively linked to a structural gene,  
CC which is then expressed at high levels.

SQ Sequence 28 BP; 6 A; 9 C; 7 G; 6 T; 0 other;

Query Match 1.4%; Score 17; DB 16; Length 28;  
Best Local Similarity 100.0%; Pred. No. 95;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 554 atccgggagaaatctc 570  
|||||

Db 5 atccgggagaaatctc 21

## RESULT 9

AAFI4771  
ID AAFI4771 standard; CDNA; 170 BP.

AC AAFI4771;

DT 13-MAR-2001 (first entry)

DE Aspergillus oryzae EST SEQ ID NO:7294.

XX Multiple gene expression; filamentous fungal cell; EST;

KM expressed sequence tag; Fusarium venenatum; Aspergillus niger;

KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;  
KW culture condition; environmental stress; spore morphogenesis;  
KW metabolic pathway engineering; catabolic pathway engineering; ss.

OS Aspergillus oryzae.

PN WO200056762-A2.

PD 28-SEP-2000.

PF 22-MAR-2000; 2000WO-US07781.

PR 22-MAR-1999; 99US-0273623.

PA (NOVO) NOVO NORDISK BIOTECH INC.

XX (NOVO) NOVO NORDISK AS.

PI Berka RM, Rey MM, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;

DR WPI; 2000-594572/56.

PT Monitoring differential expression of genes in filamentous fungal cells  
PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
PT substrate of expressed sequence tags -

PS Claim 88; Page 2950; 3161pp; English.

CC The present invention describes a method for monitoring differential  
CC expression of genes in a first filamentous fungal (FF) cell relative to  
CC expression of the same genes in one or more second filamentous fungal  
CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
CC are used in the methods for monitoring differential expression of genes  
CC in a first filamentous fungal (FF) cell relative to expression of the  
CC same genes in one or more second filamentous fungal cells. Monitoring  
CC the global expression of genes from FF cells allows the production  
CC potential of the microorganisms to be improved. New genes may be  
CC discovered, possible functions of unknown open reading frames can be  
CC identified and gene copy number variation and stability can be  
CC monitored. The expression of genes can be used to study how FF cells  
CC adapt to changes in culture conditions, environmental stress, spore  
CC morphogenesis, recombination, metabolic or catabolic pathway  
CC engineering. Using ESTs provides several advantages over genomic or  
CC random cDNA clones including elimination of redundancy as one spot on an  
CC array equals one gene or open reading frame, and organisation of the  
CC microarrays based on function of the gene products to facilitate  
CC analysis of the results. AAFI478 to AAFI1247 represents ESTs from  
CC Fusarium venenatum; AAFI1248 to AAFI1853 represents ESTs from Aspergillus  
CC niger; AAFI1854 to AAFI4878 represents ESTs from Aspergillus oryzae; and  
CC AAFI4879 to AAFI5337 represents ESTs from Trichoderma reesei, which are  
CC all specifically claimed in the present invention.

SQ Sequence 170 BP; 34 A; 48 C; 55 G; 32 T; 1 other;

Query Match 1.4%; Score 17; DB 21; Length 170;  
Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 876 tgaattccttgatcact 892  
|||||

Db 130 tgaattccttgatcact 146

## RESULT 10

AAV78693/C  
ID AAV78693 standard; DNA; 304 BP.

AC AAV78693;

DT 16-MAR-1999 (first entry)

DE Staphylococcus aureus contig SEQ ID #4382.

XX Computer readable medium; vaccine; S.aureus infection; immunodetection;  
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;

KW skin infection; surgical wound infection; scalded skin syndrome;



immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory; cytosstatic; antibacterial; antifungal; antiviral; antidiabetic; antistatic; vulnerrary; antiparkinsonian; antilicer; osteopathic; neuroprotective; nootropic; antipsoriatic; cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine; autoimmune disorder; multiple sclerosis; allergic condition; insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer; lymphoid cell deficiency; burn; osteoporosis; osteoarthritis; central nervous system disorder; Alzheimer's disease; stroke; Parkinson's disease; Huntington's disease; coagulation disorder; haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour; infection; depression; psoriasis; ss.

OS Homo sapiens.  
 PN WO200021991-A1.  
 PD 20-APR-2000.  
 PF 15-OCT-1999; 99WO-US24206.  
 PR 15-OCT-1998; 98US-0104436.  
 PA (GENY ) GENETICS INST INC.  
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C, Merberg D, Treacy M, Bowman MR;  
 DR WPI: 2000-317938/27.  
 PT Isolated polynucleotides, and encoded proteins, comprising secreted expressed sequence tags (seSTs), useful for treating various disorders such as autoimmune, infectious, and central nervous system disorders -

PS Claim 1: Page 742; 803pp; English.

XX AAA43426 to AAA45925 represent specifically claimed secreted expressed sequence tags (seSTs), isolated from human, mouse, chicken and rat tissue sources. The seSTs can have a range of activities depending on the tissues they were isolated from. The activities include:

XX chemotactic; proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory; CC cytosstatic; antibacterial; antifungal; antiviral; antidiabetic; CC antistatic; vulnerrary; antilicer; osteopathic; neuroprotective; CC nootropic; antiparkinsonian; antipsoriatic; cerebroprotective; CC anticonvulsant; and antidepressant. The seSTs can be used for gene therapy and in vaccines. The seSTs are useful as probes for the identification and isolation of full-length cDNAs and genomic DNA molecules which correspond to the seSTs. Proteins encoded by the seSTs are useful in assays for determining biological activity and raising antibodies. They may be useful for treatment of autoimmune disorders (multiple sclerosis, insulin dependent diabetes), allergic conditions (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers, CC osteoporosis, osteoarthritis, central nervous system disorders CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's disease), tumours, bacterial, fungal or viral infections, depression and CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given in the exemplification of the present invention.

XX Sequence 122 BP; 35 A; 13 C; 26 G; 46 T; 2 other;

XX SQ

Query Match 1.6%; Score 19; DB 21; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 9.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 73 aaagaaaaatacacatca 91  
 ||||||||||||||||  
 DB 80 AAAGAAAATACACATCA 62

RESULT 7

AAT72956  
 ID AAT72956 standard; DNA; 388 BP.  
 XX AAT72956;  
 AC AAT72956;  
 XX 10-MAR-1998 (first entry)  
 DT  
 DE PC-gene 2 putative eukaryotic promoter element.  
 XX  
 XX Glyceraldehyde-3-phosphate dehydrogenase; GAPDH; carotenoid;  
 KW synthesis; promoter; recombinant DNA; astaxanthin; ribosomal protein;  
 KW food colouring; ds.  
 XX  
 XX Phaffia rhodozyma.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Promoter 1..385  
 FT /tag= a  
 FT 249..263  
 FT TATA\_signal  
 FT /tag= b  
 FT /note= "putative"  
 FT 287..302  
 FT /tag= c  
 FT /function= cap-signal  
 FT /note= "putative"  
 FT 386..388  
 FT /tag= d  
 FT /function= start\_of\_CDS  
 FT 85  
 FT /tag= e  
 FT /note= "uncertain"

XX W09723633-A1.  
 PN 03-JUL-1997.  
 PD  
 PF 23-DEC-1996; 96WO-EP05887.  
 PR 11-APR-1996; 96EP-0200943.  
 PR 22-DEC-1995; 95EP-0203620.  
 XX  
 PA (KONN ) GIST-BROCADES BV.  
 PA (OOIJEN A J J.  
 PI Verdoes JC, Wery J;  
 XX WPI: 1997-351068/32.  
 DR  
 XX Phaffia derived GAPDH and carotenoid synthesis genes and promoter  
 PT fragment - used in the recombinant production of therapeutically  
 PT useful proteins e.g. carotenoids for use in food colouring

XX Example 4: Page 47; 118pp; English.

PS The present sequence represents a PC-gene 2 putative eukaryotic promoter element, which is mentioned in example 4 of the present invention describing novel recombinant DNA. The recombinant DNA comprises a transcription promoter operably linked to a downstream sequence to be expressed, where the transcription promoter comprises a region found upstream of the open reading frame (ORF) of a highly expressed Phaffia gene (preferably GAPDH, a ribosomal protein gene, an enzyme involved in the biosynthesis pathway). The recombinant DNA can be used to transform hosts, preferably Phaffia. These transformed hosts are then used in the recombinant production of GAPDH or an enzyme involved in carotenoid synthesis, preferably astaxanthin. They may also be used to produce a pharmaceutical product. Purified carotenoids can be used as colourants in food and/or feed, and also in cosmetics.

XX Sequence 388 BP; 97 A; 121 C; 77 G; 92 T; 1 other;

XX SQ

Query Match 1.5%; Score 18; DB 18; Length 388;  
 Best Local Similarity 100.0%; Pred. No. 34;

Query Match 1.9%; Score 23; DB 21; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.084;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 470 ctgcgtactctctgcgcatgt 492  
|||||  
Db 1 ctgcgtactctctgcgcatgt 23

## RESULT 4

AA294283  
ID AA294283 standard; DNA; 23 BP.

AC AA294283;

DT 03-JUL-2000 (first entry)

DE PHELIIX gene PCR primer 22P4G9.1.

XX PHELIIX; human; testis-specific; transcription factor;

KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;

KW gene therapy; diagnosis; vaccine; PCR primer;

KW chromosome 13q13.1-13.3; ss.

XX Homo sapiens.

OS WO200012709-A2.

PN 09-MAR-2000.

XX 31-AUG-1999; 99WO-US20137.

PF 31-AUG-1998; 98US-0098610.

PR 31-OCT-1998; 98US-0106524.

XX (UROG-) UROGENESIS INC.

PA (AFAR/) AFAR D E.

PA (HUBE/) HUBERT R S.

PA (RAIT/) RAITANO A B.

XX Afar DE, Hubert RS, Raitano AB;

XX WPI: 2000-237872/20.

PT Testis specific Helix Loop Helix proteins expressed in cancers and

PT useful for the prevention, diagnosis and treatment of prostate, bladder

PT and ovarian tumors -

XX Example 9; Page 38; 62pp; English.

XX The present sequence is that of PCR primer 22P4G9.1, which was used

CC for chromosomal mapping of the PHELIIX gene (see AA294275) to human

CC chromosome 13q13.1-13.3. PHELIIX (see AAY79269) is a transcription

CC factor that is normally expressed only in testis tissue, but which

CC is up-regulated in prostate and other types of cancer. The

CC invention provides diagnostic and therapeutic methods useful in the

CC management of various cancers which express PHELIIX, including

CC prostate cancer, bladder cancer, ovarian cancer and testicular

CC yncancer.

## RESULT 5

AAA54515/C  
ID AAA54515 standard; DNA; 609 BP.

AC AAA54515;

DT 11-APR-2001 (first entry)

DE Fructan exohydrolase (FEH) partial sequence.

XX Fructan exohydrolase; FEH; transgenic plant; recombination;

KW transgene; gene expression; detergent; detergent additive;

KW oral care composition; ds.

XX Triticum aestivum.

OS WO200068402-A1.

PN 16-NOV-2000.

XX 08-MAY-2000; 2000WO-EP04226.

PF 06-MAY-1999; 99BE-0000329.

PR (LEUV-) LEUVEN RES & DEV.

XX Van Den Ende W, Van Laere A, De Roover J, Michiels A;

PI WPI: 2001-007401/01.

XX Novel DNA molecules encoding enzymes having fructan exohydrolase

PT activity for use in transgenic plant production, dental care

PT compositions, and in detergents

XX Claim 11; Fig 6-1; 45pp; English.

XX Transgenic plants such as Cichorium intybus, Cynara scolymus,

CC Helianthus tuberosus, Scorzonera hispanica, Oryza sativa, Zea mays,

CC Triticum aestivum, Triticum durum, Hordeum vulgare, Secale cereale,

CC Avena sativa, Sorghum vulgare, Phleum pratense, Lolium temulentum,

CC Dactylis glomerata, Pennisetum americanum, Allium cepa, Agave

CC americanum, Agave azul tequilana, Sorghum bicolor and Panicum miliaceum,

CC transformed with a vector encoding a fructan exohydrolase (FEH)

CC enzyme are useful for the recombinant production of FEH or other

CC polypeptides having FEH activity. The FEH polypeptides produced

CC are useful in detergents or as a detergent additive and in oral

CC care compositions.

XX Sequence 609 BP; 159 A; 121 C; 164 G; 165 T; 0 other;

XX Query Match 1.6%; Score 20; DB 22; Length 609;

Best Local Similarity 100.0%; Pred. No. 3.3;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1148 caaagaatgatttgaagc 1167  
|||||  
Db 119 CAAAGATGATTTTGAAGC 100

## RESULT 6

AAA45601/C  
ID AAA45601 standard; CDNA; 122 BP.

AC AAA45601;

DT 21-AUG-2000 (first entry)

DE Human secreted expressed sequence tag SEQ ID NO:2176.

XX Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;

KW expressed sequence tag; EST; probe; chemotactic; proliferative;

PT Testis specific Helix Loop Helix proteins expressed in cancers and  
PT useful for the prevention, diagnosis and treatment of prostate, bladder  
PT and ovarian tumors -  
XX  
XX  
PS Example 1: Page 34; 62pp; English.  
XX  
CC The present sequence is that of a primer used in RT-PCR analysis  
CC to determine expression levels of the human PHELIIX gene (see also  
CC AA294275). Examination of 1st strand cDNAs from 16 normal human  
CC tissues detected expression of the PHELIIX gene only in testis  
CC tissue. PHELIIX expression was also detected in LAPC-4 xenografts.  
CC PHELIIX (see AAY79269) is a transcription factor that is normally  
CC expressed only in testis tissue, but which is up-regulated in  
CC prostate and other types of cancer. The invention provides  
CC diagnostic and therapeutic methods useful in the management of  
CC various cancers which express PHELIIX, including prostate cancer,  
CC bladder cancer, ovarian cancer and testicular cancer.  
CC  
XX  
SQ Sequence 24 BP; 2 A; 3 C; 7 G; 12 T; 0 other;

Query Match 2.0%; Score 24; DB 21; Length 24;  
Best Local Similarity 100.0%; Pred. No. 0.026;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 627 agaacaacaacaccattgagc 650  
DB 24 AGAACAACAACACCATTGAGC 1

RESULT 2  
AA294284/C  
ID AA294284 standard; DNA; 24 BP.  
XX  
XX AA294284;  
XX  
XX 03-JUL-2000 (first entry)  
XX  
XX  
DE PHELIIX gene PCR primer 22P4G9.2.  
XX  
XX  
XX PHELIIX, human; testis-specific; transcription factor;  
XX prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
XX gene therapy; diagnosis; vaccine; PCR primer;  
XX chromosome 13q13.1-13.3; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200012709-A2.  
XX  
XX PD 09-MAR-2000.  
XX  
XX PF 31-AUG-1999; 99WO-US20137.  
XX  
XX PR 31-AUG-1998; 98US-0098610.  
XX PR 31-OCT-1998; 98US-0106524.  
XX  
XX (UROC-) UROGENESYS INC.  
XX (AFAR/) AFAR D E.  
XX (HUBE/) HUBERT R S.  
XX (RAIT/) RAITANO A B.  
XX  
XX PI Afar DE, Hubert RS, Raitano AB;  
XX  
XX DR WPI; 2000-237872/20.  
XX  
XX PT Testis specific Helix Loop Helix proteins expressed in cancers and  
XX useful for the prevention, diagnosis and treatment of prostate, bladder  
XX and ovarian tumors -  
XX  
XX PS Example 9: Page 38; 62pp; English.  
XX  
XX The present sequence is that of PCR primer 22P4G9.2, which was used  
CC for chromosomal mapping of the PHELIIX gene (see AA294275) to human

CC chromosome 13q.13.1-13.3. PHELIIX (see AAY79269) is a transcription  
CC factor that is normally expressed only in testis tissue, but which  
CC is up-regulated in prostate and other types of cancer. The  
CC invention provides diagnostic and therapeutic methods useful in the  
CC management of various cancers which express PHELIIX, including  
CC prostate cancer, bladder cancer, ovarian cancer and testicular  
CC cancer.  
XX  
XX  
SQ Sequence 24 BP; 2 A; 3 C; 7 G; 12 T; 0 other;

Query Match 2.0%; Score 24; DB 21; Length 24;  
Best Local Similarity 100.0%; Pred. No. 0.026;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 627 agaacaacaacaccattgagc 650  
DB 24 AGAACAACAACACCATTGAGC 1

RESULT 3  
AA294281  
ID AA294281 standard; DNA; 23 BP.  
XX  
XX AA294281;  
XX  
XX 03-JUL-2000 (first entry)  
XX  
XX  
DE PHELIIX gene PCR primer.  
XX  
XX  
XX PHELIIX, human; testis-specific; transcription factor;  
XX prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
XX gene therapy; diagnosis; vaccine; PCR primer; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200012709-A2.  
XX  
XX PD 09-MAR-2000.  
XX  
XX PF 31-AUG-1999; 99WO-US20137.  
XX  
XX PR 31-AUG-1998; 98US-0098610.  
XX PR 31-OCT-1998; 98US-0106524.  
XX  
XX (UROC-) UROGENESYS INC.  
XX (AFAR/) AFAR D E.  
XX (HUBE/) HUBERT R S.  
XX (RAIT/) RAITANO A B.  
XX  
XX PI Afar DE, Hubert RS, Raitano AB;  
XX  
XX DR WPI; 2000-237872/20.  
XX  
XX PT Testis specific Helix Loop Helix proteins expressed in cancers and  
XX useful for the prevention, diagnosis and treatment of prostate, bladder  
XX and ovarian tumors -  
XX  
XX PS Example 1: Page 34; 62pp; English.  
XX  
XX The present sequence is that of a primer used in RT-PCR analysis  
XX to determine expression levels of the human PHELIIX gene (see also  
XX AA294275). Examination of 1st strand cDNAs from 16 normal human  
XX tissues detected expression of the PHELIIX gene only in testis  
XX tissue. PHELIIX expression was also detected in LAPC-4 xenografts.  
XX PHELIIX (see AAY79269) is a transcription factor that is normally  
XX expressed only in testis tissue, but which is up-regulated in  
XX prostate and other types of cancer. The invention provides  
XX diagnostic and therapeutic methods useful in the management of  
XX various cancers which express PHELIIX, including prostate cancer,  
XX bladder cancer, ovarian cancer and testicular cancer.  
XX  
XX  
SQ Sequence 23 BP; 2 A; 7 C; 5 G; 9 T; 0 other;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2001, 11:36:05 : Search time 126.73 Seconds  
(without alignments)  
6019.888 Million cell updates/sec

Title: US-09-389-000-1\_COPY\_735\_1949  
Perfect score: 1215  
Sequence: 1 ggctctctgaaggtgcttctt.....ttgcatggttgcctctc 1215

Scoring table: OLIGO\_MNC  
Gapop 60.0 , Gapext 60.0

Searched: 730101 seqs, 313950809 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1332894

Minimum DB seq length: 0

Maximum DB seq length: 1214

Post-processing: Listing first 45 summaries

Database:

N\_Geneseq\_0601:\*

1: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1980.DAT:\*

2: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1981.DAT:\*

3: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1982.DAT:\*

4: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1983.DAT:\*

5: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1984.DAT:\*

6: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1985.DAT:\*

7: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1986.DAT:\*

8: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1987.DAT:\*

9: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1988.DAT:\*

10: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1989.DAT:\*

11: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1990.DAT:\*

12: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1991.DAT:\*

13: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1992.DAT:\*

14: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1993.DAT:\*

15: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1994.DAT:\*

16: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1995.DAT:\*

17: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1996.DAT:\*

18: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1997.DAT:\*

19: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1998.DAT:\*

20: /cgnl\_9/gcgdata/geneseq/geneseqn/NA1999.DAT:\*

21: /cgnl\_9/gcgdata/geneseq/geneseqn/NA2000.DAT:\*

22: /cgnl\_9/gcgdata/geneseq/geneseqn/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	24	2.0	24	21	AAZ94282
C 2	24	2.0	24	21	AAZ94284
C 3	23	1.9	23	21	AAZ94281
C 4	23	1.9	23	21	AAZ94283
C 5	20	1.6	609	22	AA454515
C 6	19	1.6	122	21	AAA45601
C 7	18	1.5	388	18	AAAT72956
C 8	17	1.4	28	16	AAO86825
C 9	17	1.4	170	21	AAFI4771
C 10	17	1.4	304	18	AAV78693
C 11	17	1.4	330	21	AAFI1660

12	17	1.4	341	21	AAC28532
C 13	17	1.4	344	21	AAC21590
C 14	17	1.4	370	18	AAV78456
C 15	17	1.4	400	18	AAV78096
C 16	17	1.4	403	22	AAE93376
C 17	17	1.4	457	21	AAA28157
C 18	17	1.4	495	21	AAC01443
C 19	17	1.4	597	21	AAC03992
C 20	17	1.4	721	21	AAC33953
C 21	17	1.4	1053	18	AAE83926
C 22	16	1.3	17	20	AAAI7522
C 23	16	1.3	17	20	AAAI7523
C 24	16	1.3	115	20	AAV71601
C 25	16	1.3	115	21	AAE23757
C 26	16	1.3	178	16	AAE24565
C 27	16	1.3	187	21	AAC25075
C 28	16	1.3	236	21	AAC25366
C 29	16	1.3	239	20	AAV89201
C 30	16	1.3	241	16	AAE20599
C 31	16	1.3	285	21	AAC04945
C 32	16	1.3	289	21	AAC18840
C 33	16	1.3	291	21	AAC28811
C 34	16	1.3	292	21	AAZ42513
C 35	16	1.3	300	20	AAZ13919
C 36	16	1.3	300	20	AAZ13489
C 37	16	1.3	315	19	AAV20144
C 38	16	1.3	325	16	AAE22640
C 39	16	1.3	335	21	AAAB1995
C 40	16	1.3	404	21	AAC01745
C 41	16	1.3	406	21	AAC94514
C 42	16	1.3	408	14	AAO59868
C 43	16	1.3	414	21	AAA31492
C 44	16	1.3	452	21	AAA45874
C 45	16	1.3	457	21	AAC19276

## ALIGNMENTS

RESULT 1	
ID AAZ94282 standard; DNA; 24 BP.	
XX AC AAZ94282;	
XX AC 03-JUL-2000 (first entry)	
DT	
XX	
DE PHELI gene PCR primer.	
XX	
KW PHELI gene; human; testis-specific; transcription factor;	
KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;	
KW gene therapy; diagnosis; vaccine; PCR primer; ss.	
XX	
OS Homo sapiens.	
XX	
PN W0200012709-A2.	
XX	
PD 09-MAR-2000.	
XX	
PF 31-AUG-1999; 99WO-US20137.	
XX	
PR 31-AUG-1998; 98US-0098610.	
PR 31-OCT-1998; 98US-0106524.	
XX	
PA (UROC-) UROGENESYS INC.	
PA (AFAR/) AFAR D E.	
PA (HUBE/) HUBERT R S.	
PA (RAIT/) RAITANO A B.	
XX	
PI Afar DE, Hubert RS, Raitano AB;	
XX	
DR WPI: 2000-237872/20.	
XX	

Human secreted pro  
Human secreted pro  
Staphylococcus aur  
Staphylococcus aur  
CDNA encoding SRT  
Human purh amplico  
Human secreted pro  
Human secreted pro  
Arabidopsis thalia  
DNA encoding a Sta  
Aryl hydrocarbon n  
Aryl hydrocarbon n  
Bacillus subtilis  
55 mer oligonucleo  
Human gene signatu  
Human secreted pro  
Human secreted pro  
EST clone C06. Ho  
Human gene signatu  
Human secreted pro  
Human secreted pro  
Human 5' EST Isola  
Human gene express  
Human gene express  
Probe (45) for mic  
Human gene signatu  
N. meningitidis pa  
Human secreted pro  
Cat flea hindgut a  
Human brain expres  
Plant microsatelli  
Human secreted exp  
Human secreted pro

**THIS PAGE BLANK (USPTO)**



http://www.bhmp.mrc.ac.uk/ for details  
or contact: biohelp@bhmp.mrc.ac.uk.

## FEATURES

## SOURCE

Location/Qualifiers

1. 177

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/sex="male"

/dev\_stage="adult"

/tissue\_type="blood"

/clone\_lib="CGI-1"

/clone="36d4"

## BASE COUNT

60 a 32 c 32 g 53 t

## ORIGIN

Query Match 1.4%; Score 17; DB 92; Length 177;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 940 aacaatcagacattt 956

DB 161 AACAAATCAGACATT 177

## RESULT 15

## HS36D4F/c

## LOCUS

## DEFINITION

H.sapiens Cpg island DNA genomic MseI fragment, clone 36d4, forward

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

human.

Search completed: September 15, 2001, 12:12:19  
Job time: 5911 sec

Vector: pGEM-5Zf(-)  
Clones are available from the UK MRC Human Genome Mapping Project  
Resource Centre, Hinxton, Cambridgeshire CB10 1RO, UK. See URL:  
http://www.bhmp.mrc.ac.uk/ for details  
or contact: biohelp@bhmp.mrc.ac.uk.

## FEATURES

## SOURCE

1. 179

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/sex="male"

/dev\_stage="adult"

/tissue\_type="blood"

/clone\_lib="CGI-1"

/clone="36d4"

## BASE COUNT

53 a 32 c 34 g 59 t 1 others

## ORIGIN

Query Match 1.4%; Score 17; DB 92; Length 179;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 940 aacaatcagacattt 956

DB 18 AACAAATCAGACATT 2

ACCESSION	Candida tropicalis, sequence tagged site.
VERSION	AL439011
KEYWORDS	AT439011.1 GI:1222424
SOURCE	STS.
ORGANISM	Candida tropicalis.
REFERENCE	Candida tropicalis
AUTHORS	Eunhyota, Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
	Saccharomycetales; mitosporic Saccharomycetales; Candida.
	1 (bases 1 to 1066)
	Blandin,G., Ozier-Kalogeropoulos,O., Wincker,P., Artiguenave,F. and
	Dujon,B.
	Genomic Exploration of the Hemiascomycetous Yeasts: 16. Candida
	tropicalis
TITLE	FEMS Lett. 487 (1), 91-94 (2000)
JOURNAL	1152891
PUBMED	2 (bases 1 to 1066)
REFERENCES	Souci�t,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
	Boletin-Fukuhara,M., Bon,E., Brotier,P., Casaregola,S.,
	de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Lorente,B.,
	Maiertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
	Saurin,W., Tekala,F., Toffano-Nic�che,C., Wesolowski-Louvel,M.,
	Wincker,P. and Weissenbach,J.
	Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of
	yeast species for molecular evolution studies(1)
	FEMS Lett. 487 (1), 3-12 (2000)
TITLE	1152876
JOURNAL	3 (bases 1 to 1066)
PUBMED	Genoscope.
REFERENCES	Direct Submission
AUTHORS	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
JOURNAL	2 rue Gaston Cremlieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
	seque@genoscope.cns.fr - Web :
	This STS is part of a random genomic sequencing program of thirteen
	yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
	exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
	Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
	lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
	angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
	Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
	5 kb were prepared and both extremities were sequenced. See
	keywords for description of this sequence and for the sequence of
	the other extremity of this insert.
FEATURES	Location/Qualifiers
SOURCE	1..1066
	/organism="Candida tropicalis"
	/strain="CBS_94"
	/db_xref="taxon:5482"
	/clone="BD0AA01H05"
	/clone_lib="BD0AA"
	/note="end : T3"
	1..1066
BASE COUNT	454 a 179 c 202 g 230 t 1 others
ORIGIN	
OY	72 aaagaataatcacatt 89
Db	
	Query Match 1.5%; Score 18; DB 53; Length 1066;
	Best Local Similarity 100.0%; Pred. No. 3.7e+02;
	Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT	13
ID	E08984 standard; DNA; UNC; 28 BP.
AC	E08984;
XX	
SV	E08984.1
OT	07-OCT-1997 (Rel. 52, Created)

DT	02-SEP-2000 (Rel. 65, Last updated, Version 2)
XX	
DE	PCR primer.
XX	
KW	JP 1995087977-A/3.
XX	
OS	unidentified
OC	unclassified.
XX	
RN	[1]
RA	1-28
RA	Inui M., Kobayashi M., Kurusu Y., Yugawa H.;
RL	"DNA FRAGMENT HAVING PROMOTER FUNCTION IN CORVINE TYPE BACTERIUM";
RL	Patent number Jp1995087977-A/3, 04-APR-1995.
RL	MITSUBISHI CHEM CORP.
XX	
OS	None
CC	Artificial sequences.
CC	OC
CC	PN JP 1995087977-A/3
CC	PD 04-APR-1995
CC	PF 12-AUG-1993 JP 1993200707
CC	PI INUI MASAYUKI, KOBAYASHI MIKI, KURUSU YASUROU, YUGAWA HIDEAKI
CC	PC C12N15/09, C12N15/09, C12R1:13;
CC	CC strandedness: Single;
CC	CC topology: Linear;
CC	CC FH key Location/Qualifiers
CC	CC FH source 1..28
CC	CC FT /Organism="Artificial sequences"
CC	CC FT
XX	
FH	Key Location/Qualifiers
FH	FT source 1..28
FT	/db_xref="taxon:32644"
FT	/organism="unidentified"
XX	
SO	Sequence 28 BP; 6 A; 9 C; 7 G; 6 T; 0 other;
QY	554 atccggagagaatctc 570
Db	5 ATCCGCGAGAAAATCTC 21
RESULT 14	
LOCUS	HS36DAR
DEFINITION	HS36DAR 177 bp DNA PRI 17-OCT-1995
ACCESSION	H.sapiens Cpg Island DNA genomic Msel fragment, clone 36d4, reverse
VERSION	read cpg36d4.rta.
KEYWORDS	Z55360
SOURCE	Z55360.1 GI:1021401
ORGANISM	Cpg Island; genomic Msel fragment.
	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 177)
AUTHORS	Macdonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE	Direct Submission
JOURNAL	Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
	CB10 IRO, England. E-mail contact: humgen@esanger.ac.uk
REFERENCE	2 (bases 1 to 177)
AUTHORS	Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE	Purification of Cpg Islands using a methylated DNA binding column
JOURNAL	Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE	94282070
COMMENT	Vector: pGEM-52f(-)
	Clones are available from the UK MRC Human Genome Mapping Project
	Resource Centre, Hinxton, Cambridgeshire CB10 IRO, UK. See URL:



```

misc_feature      <84..>434  

                  <note>"similar to Saccharomyces cerevisiae ORF YGR202c [VMA7; H+-ATPase VI domain 14 kDa subunit, vacuolar ]"  

                  /evidence=not_experimental  

BASE COUNT      269 a      161 c      127 g      349 t      2 others  

ORIGIN  

Query Match      1.5%: Score 18; DB 54; Length 908;  

Best Local Similarity 100.0%: Pred. No. 3.7e+02;  

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  

Oy      72 aaagaataacataccatt 89  

          |||||  

Db      632 AAAAGAAAAATACACATT 649  

  

RESULT 10  

CNS07BDY/C      CNS07BDY      1015 bp      DNA      STS      12-JAN-2001  

LOCUS      T3 end of clone BC0AA011D03 of library BC0AA from strain CBS 767 of  

DEFINITION      Debaryomyces hansenii, sequence tagged site.  

ACCESSION      A1437756  

VERSION      A1437756.1 GI:12221169  

KEYWORDS      STS.  

SOURCE      Debaryomyces hansenii.  

ORGANISM      Debaryomyces hansenii  

              Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  

              Saccharomycetales; Saccharomycetaceae; Debaryomyces.  

REFERENCE      1 (bases 1 to 1015)  

AUTHORS      Lepingle,A., Casaregola,S., Neuveglise,C., Bon,E., Nguyen,H.,  

              Artiguenave,F., Wincker,P. and Gaillardin,C.  

TITLE      Genomic Exploration of the Hemiascomycetous Yeasts: 14.  

JOURNAL      Debaryomyces hansenii var. hansenii  

PUBMED      FEBS Lett. 487 (1), 82-86 (2000)  

AUTHORS      1152889  

              2 (bases 1 to 1015)  

              Souciet,U.L., Aigle,M., Artiguenave,F., Blandin,G.,  

              Boletín-Pukuhara,M., Bon,E., Brothier,P., Casaregola,S.,  

              de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,  

              Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,  

              Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,  

              Wincker,P. and Weissenbach,J.  

              Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of  

              yeast species for molecular evolution studies(1)  

JOURNAL      FEBS Lett. 487 (1), 3-12 (2000)  

PUBMED      1152876  

AUTHORS      3 (bases 1 to 1015)  

              Genoscope.  

              Direct Submission  

              Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,  

              2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail :  

              seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)  

              This STS is part of a random genomic sequencing program of thirteen  

              yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces  

              exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,  

              Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces  

              lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia  

              angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,  

              Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to  

              5 kb were prepared and both extremities were sequenced. See  

              keywords for description of this sequence and for the sequence of  

              the other extremity of this insert.  

              Location/Qualifiers  

              1..1015  

              /organism="Debaryomyces hansenii"  

              /strain="CBS 767"  

              /variety="hansenii"  

              /db_xref="taxon:4959"  

              /clone="BC0AA011D03"  

              /clone_lib="BC0AA"  

              /note="end : T3"  

              <1..>986  


```

Query Match	Best Local Similarity	Score	DB	Length
Matches 18; Conservative	100.0%; Pred. No. 3,7e+02;	1015;		
OY 403 gatctctcttcttcattc 420	1.5%; Score 18; DB 53; Length 1015;			
Db 642 GATCTCTCTCTTCATTC 625				
RESULT 11				
AF115672	1019 bp	DNA	PLN	10-APR-2001
LOCUS AF115672/c				
DEFINITION	Sempervivum italicum maturase K-like protein (matk) gene, partial			
ACCESSION	AF115672			
VERSION	AF115672.1 GI:13568591			
KEYWORDS	Sempervivum italicum.			
SOURCE	Chloroplast Sempervivum italicum			
ORGANISM	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Saxifragales; Crassulaceae; Sempervivum.			
REFERENCE	1 (bases 1 to 1019)			
AUTHORS	Mort,M.E., Solitis,D.E., Solitis,P.S., Francisco-Ortega,J. and Santos-Guerra,A.			
TITLE	Phylogenetic relationships and evolution of Crassulaceae inferred from matk sequence data			
JOURNAL	Am J Bot 88 (1), 76-91 (2001)			
PUBMED	11159129			
REFERENCE	2 (bases 1 to 1019)			
AUTHORS	Mort,M.E., Solitis,D.E., Solitis,P.S., Francisco-Ortega,J. and Santos-Guerra,A.			
TITLE	Direct Submission			
JOURNAL	Submitted (21-DEC-1998) Dept. of Biological Sciences, Eastern Illinois University, Biological Sciences, Charleston, IL 61920, USA			
COMMENT	NCBI staff are still waiting for submitters to provide appropriate coding region information.			
FEATURES	Location/Qualifiers			
source	1..1019			
	/organism="Sempervivum italicum"			
	/organelle="plastid;chloroplast"			
misc_feature	/db_xref="taxon:91153"			
	1..>1019			
gene	/gene="matk"			
	/note="similar to maturase K"			
	1..>1019			
	/gene="matk"			
BASE COUNT	313 a	180 c	151 g	374 t
ORIGIN	1 others			
Query Match	Best Local Similarity	Score	DB	Length
Matches 18; Conservative	100.0%; Pred. No. 3,7e+02;	1015;		
OY 173 acaatggaanaaatgagt 190	1.5%; Score 18; DB 12; Length 1019;			
Db 1014 ACAATGGAANAATGAGT 997				
RESULT 12				
CNS07/CCT/c	1066 bp	DNA	STS	12-JAN-2001
LOCUS CNS07/CCT/c				
DEFINITION	T3 end of clone BD0AA001H05 of library BD0AA from strain CHS 94 of			

BAC ends sequenced at TIGR from the RPIII BAC library. Designed and developed at the Stanford Human Genome Center.

FEATURES  
source  
1..442  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="14"  
/clone\_lib="Human"

STS  
primer\_bind 91..113  
primer\_bind complement(346..368)  
BASE COUNT 141 a 74 c 116 g 111 t  
ORIGIN

Query Match 1.5%; Score 18; DB 54; Length 442;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 435 taagaaggaagaatca 452  
|||||  
Db 125 TAAGAAGGAAGAATCA 142

RESULT 8  
CNS06WRM 882 bp DNA STS 11-JAN-2001  
LOCUS T7 end of clone AX0AA034G03 of library AX0AA from strain CBS 7064  
DEFINITION of *Pichia farinosa*, sequence tagged site.  
ACCESSION AL418818.1 GI:12201933  
VERSION AL418818.1  
KEYWORDS STS.  
SOURCE *Pichia farinosa*.  
ORGANISM *Pichia farinosa*.  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; *Pichia*.

REFERENCE  
AUTHORS 1 (bases 1 to 882)  
de Montigny,J., Spehner,C., Souclet,J., Tekala,F., Dujon,B.,  
Wincker,P., Artiguenave,F. and Potier,S.  
Genomic Exploration of the Hemiascomycetous Yeasts: 15. *Pichia*  
*sorbitophila*

JOURNAL  
PUBMED 1152890  
REFERENCE  
AUTHORS 2 (bases 1 to 882)  
Souclet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,  
Bojotin-Fukuhara,M., Bon,E., Brotlier,P., Casaregola,S.,  
de Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,  
Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,  
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,  
Wincker,P. and Weissenbach,J.  
Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of  
yeast species for molecular evolution studies(1)

TITLE  
JOURNAL  
PUBMED 1152876  
REFERENCE  
AUTHORS 3 (bases 1 to 882)  
Genoscope.  
Direct Submission  
Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,  
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
This STS is part of a random genomic sequencing program of thirteen  
yeast species: *Saccharomyces bayanus* var. *uvatum*, *Saccharomyces*  
*exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*,  
*Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces*  
*lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia*  
*angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,  
*Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to  
5 kb were prepared and both extremities were sequenced. See  
keywords for description of this sequence and for the sequence of  
the other extremity of this insert.

FEATURES  
source  
1..882  
/organism="Pichia farinosa"  
Location/Qualifiers

/strain="CBS 7064"  
/db\_xref="taxon:4920"  
/clone="AX0AA034G03"  
/clone\_lib="AX0AA"  
/note="end : T7"  
STS  
BASE COUNT 318 a 161 c 149 g 208 t 46 others  
ORIGIN

Query Match 1.5%; Score 18; DB 53; Length 882;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1138 aaaagtggacaagaat 1155  
|||||  
Db 14 AAAAGTGGACAAGAAT 31

RESULT 9  
CNS07DTH 908 bp DNA STS 12-JAN-2001  
LOCUS T7 end of clone BD0AA015E12 of library BD0AA from strain CBS 94 of  
DEFINITION *Candida tropicalis*, sequence tagged site.  
ACCESSION AL440907.1 GI:12224318  
VERSION AL440907.1  
KEYWORDS STS.  
SOURCE *Candida tropicalis*.  
ORGANISM *Candida tropicalis*.  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; mitosporic Saccharomycetales; *Candida*.  
Dujon,B.

REFERENCE  
AUTHORS 1 (bases 1 to 908)  
Blandin,G., Ozier-Kalogeropoulos,O., Wincker,P., Artiguenave,F. and  
Dujon,B.  
Genomic Exploration of the Hemiascomycetous Yeasts: 16. *Candida*  
*tropicalis*

JOURNAL  
PUBMED 1152891  
REFERENCE  
AUTHORS 2 (bases 1 to 908)  
Souclet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,  
Bojotin-Fukuhara,M., Bon,E., Brotlier,P., Casaregola,S.,  
de Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,  
Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,  
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,  
Wincker,P. and Weissenbach,J.  
Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of  
yeast species for molecular evolution studies(1)

TITLE  
JOURNAL  
PUBMED 1152876  
REFERENCE  
AUTHORS 3 (bases 1 to 908)  
Genoscope.  
Direct Submission  
Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,  
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
segref@genoscope.cns.fr - Web :  
www.genoscope.cns.fr)  
This STS is part of a random genomic sequencing program of thirteen  
yeast species: *Saccharomyces bayanus* var. *uvatum*, *Saccharomyces*  
*exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*,  
*Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces*  
*lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia*  
*angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,  
*Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to  
5 kb were prepared and both extremities were sequenced. See  
keywords for description of this sequence and for the sequence of  
the other extremity of this insert.

COMMENT  
JOURNAL  
PUBMED 1152876  
REFERENCE  
AUTHORS 3 (bases 1 to 908)  
Genoscope.  
Direct Submission  
Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,  
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
segref@genoscope.cns.fr - Web :  
www.genoscope.cns.fr)  
This STS is part of a random genomic sequencing program of thirteen  
yeast species: *Saccharomyces bayanus* var. *uvatum*, *Saccharomyces*  
*exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*,  
*Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces*  
*lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia*  
*angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,  
*Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to  
5 kb were prepared and both extremities were sequenced. See  
keywords for description of this sequence and for the sequence of  
the other extremity of this insert.

FEATURES  
source  
1..908  
/organism="Candida tropicalis"  
/strain="CBS 94"  
/db\_xref="taxon:5482"  
/clone="BD0AA015E12"  
/clone\_lib="BD0AA"  
/note="end : T7"  
Location/Qualifiers

Polymerization: 72 degrees C for 30 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9600

Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Tag polymerase: 0.05 units/ul  
Total Vol: 10 ul

Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 20 mM  
pH: 8.3

Plasmid clones, generated from a lymphoblastoid cell line from a human male, localized to human chromosome 19 by analysis on the NIGMS Human/Rodent Somatic Cell Hybrid Panel #1, Coriell Institute for Medical Research, Camden, NJ 08103.

Location/Qualifiers  
1..400  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="19"  
/clone="pg-4539"  
/clone\_lib="Human"  
STS  
Primer\_bind 215..380  
Primer\_bind complement(360..380)  
BASE COUNT 133 a 70 c 77 g 120 t  
ORIGIN

Query Match 1.5%; Score 18; DB 54; Length 400;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

615 tgaagtttgaagaac 632  
|||||  
329 tgaagtttgaagaac 312

RESULT 6  
AB04300510 429 bp DNA ROD 30-MAR-2001  
LOCUS  
DEFINITION Mus musculus oalp2 gene for organic anion transporting polypeptide 2, exon 11.  
ACCESSION AB043018  
VERSION AB043018.1 GI:11275560  
KEYWORDS  
SEGMENT  
SOURCE  
10 of 15  
Mus musculus (strain:129/SVJ) cell\_line:RM4 DNA, clone\_lib:Genome systems BAC ES(1).  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (sites)  
Ogura, K., Choudhuri, S. and Klaassen, C.D.  
Mus musculus organic anion transporting polypeptide 2 gene  
Published Only in Database (2000) In press  
2 (sites)  
Ogura, K., Choudhuri, S. and Klaassen, C.D.  
Genomic organization and tissue-specific expression of splice variants of mouse organic anion transporting polypeptide 2  
Biochem. Biophys. Res. Commun. 281 (2), 431-439 (2001)  
21092843  
3 (bases 1 to 429)  
Ogura, K., Choudhuri, S. and Klaassen, C.D.  
Direct Submission  
Submitted (19-MAY-2000) Kenichiro Ogura, University of Kansas Medical Center, Pharmacology, 3901 Rainbow Blvd., Kansas City, KS 66160, USA (E-mail:ogurakeps.royaku.ac.jp, Tel:1-913-588-7715,

FEATURES  
source  
Location/Qualifiers  
1..429  
/organism="Mus musculus"  
/strain="129/SVJ"  
/db\_xref="taxon:10090"  
/cell\_line="RM4"  
/clone\_lib="Genome systems BAC ES(1)"  
58..222  
/gene="oalp2"  
/note="CDS is reported in Acc# AB043023"  
/number=11  
/evidence=experimental  
/product="organic anion transporting polypeptide 2"

BASE COUNT 120 a 92 c 74 g 143 t  
ORIGIN

Query Match 1.5%; Score 18; DB 94; Length 429;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

956 tcaatcattacctca 973  
|||||  
145 tcaatcattacctca 162

RESULT 7  
G51525 442 bp DNA STS 30-MAR-2000  
LOCUS  
DEFINITION SHGC-80856 Human Homo sapiens STS genomic, sequence tagged site.  
ACCESSION G51525  
VERSION G51525.1 GI:5222702  
KEYWORDS  
STS.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 442)  
Olivier, M. and Cox, D.R.  
Unpublished, Olivier, M., Cox, D.R. (2000)  
Unpublished (2000)  
COMMENT  
Contact: Michael Olivier, David R. Cox  
Stanford Human Genome Center  
Stanford University School of Medicine  
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA  
Tel: (650) 320-5800  
Fax: (650) 320-5801  
E-mail: olivier@shgc.stanford.edu  
Primer A: GCCATGTCACAGCTAAATATCAG  
Primer B: GAAGCTGCAACTGCTCAAAAT  
STS size: 278  
PCR Profile:  
Initial incubation: 95 degrees C for 10 minutes  
Denaturation: 94 degrees C for 30 seconds  
Annealing: 60 degrees C for 30 seconds  
Polymerization: 72 degrees C for 23 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9700

Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Amplifit Gold Polymerase: 0.07 units/ul  
Total Vol: 5 ul

Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 8.3

FEATURES  
source  
Location/Qualifiers  
1..429  
/organism="Mus musculus"  
/strain="129/SVJ"  
/db\_xref="taxon:10090"  
/cell\_line="RM4"  
/clone\_lib="Genome systems BAC ES(1)"  
58..222  
/gene="oalp2"  
/note="CDS is reported in Acc# AB043023"  
/number=11  
/evidence=experimental  
/product="organic anion transporting polypeptide 2"

BASE COUNT 120 a 92 c 74 g 143 t  
ORIGIN

Query Match 1.5%; Score 18; DB 94; Length 429;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

956 tcaatcattacctca 973  
|||||  
145 tcaatcattacctca 162

RESULT 7  
G51525 442 bp DNA STS 30-MAR-2000  
LOCUS  
DEFINITION SHGC-80856 Human Homo sapiens STS genomic, sequence tagged site.  
ACCESSION G51525  
VERSION G51525.1 GI:5222702  
KEYWORDS  
STS.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 442)  
Olivier, M. and Cox, D.R.  
Unpublished, Olivier, M., Cox, D.R. (2000)  
Unpublished (2000)  
COMMENT  
Contact: Michael Olivier, David R. Cox  
Stanford Human Genome Center  
Stanford University School of Medicine  
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA  
Tel: (650) 320-5800  
Fax: (650) 320-5801  
E-mail: olivier@shgc.stanford.edu  
Primer A: GCCATGTCACAGCTAAATATCAG  
Primer B: GAAGCTGCAACTGCTCAAAAT  
STS size: 278  
PCR Profile:  
Initial incubation: 95 degrees C for 10 minutes  
Denaturation: 94 degrees C for 30 seconds  
Annealing: 60 degrees C for 30 seconds  
Polymerization: 72 degrees C for 23 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9700

Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Amplifit Gold Polymerase: 0.07 units/ul  
Total Vol: 5 ul

Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 8.3

```

RESULT 3
LOCUS XCU70906 1167 bp DNA
DEFINITION Xestia c-nigrum granulosis virus AcNPV ORF78 homolog gene, complete
            cds, and glycoprotein 41 homolog gene, partial cds.
ACCESSION U70906
VERSION U70906.1 GI:3033361
KEYWORDS
SOURCE Xestia c-nigrum granulovirus.
ORGANISM Xestia c-nigrum granulovirus
REFERENCE 1 (bases 1 to 1167)
AUTHORS Goto, C., Hayakawa, T. and Maeda, S.
TITLE Genome organization of Xestia c-nigrum granulovirus
JOURNAL Virus Genes 16 (2), 199-210 (1998)
MEDLINE 98271593
PUBMED 9608666
REFERENCE 2 (bases 1 to 1167)
AUTHORS Maeda, S.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1996) Department of Entomology, University of
            California, Davis, CA 95616, USA
REFERENCE 3 (bases 1 to 1167)
AUTHORS Maeda, S.
TITLE Direct Submission
JOURNAL Submitted (07-APR-1998) Department of Entomology, University of
            California, Davis, CA 95616, USA
REMARK Sequence update by submitter
COMMENT On Apr 7, 1998 this sequence version replaced gi:1835321.
FEATURES
            location/Qualifiers
                1..1167
                /organism="Xestia c-nigrum granulovirus"
                /strain="alpha-4"
                /db_xref="taxon:51677"
                /clone="C006"
                <1..726
                /note="gp41: similar to sequence with Swiss-Prot Accession
                Number P32651"
                /codon_start=1
                /product="glycoprotein 41 homolog"
                /protein_id="AAC12813.1"
                /db_xref="GI:3033362"
                /translation="TKRRETDQLEYYKNEIKYQGNNGKNGPHGIFDPSFYSPI
                MKSTADKFKRRLNLAHLSNVKYOATATITONKPUPLVONCDYADEFKLCHKAN
                VSTNIKOLNERTNSRLNSCTNLNMLVDVLTGAHEGYNNCLNDMKAKYKFERD
                DVSEPLVQAPLNMSTNVFALIDAAKKYKGPATEVYADHVKRLNSQTASTSOLYATEL
                AFENNALRGLIQOLNSYETLS"
                775..779
                /note="late promoter"
                790..1101
                /note="similar to sequence with Swiss-Prot Accession
                Number Q06693"
                /codon_start=1
                /product="AcNPV ORF78 homolog"
                /protein_id="AAC12814.1"
                /db_xref="GI:1835323"
                /translation="MQQHLIDIPFDRLTIPDAVDAIPLKLAYSKSSDDNKPYPVSQAQ
                AVYGSREKSAQSDMSNMVFALACITVLVYIMLSYIVSLRTNNAPLRDYEDDEF
                E"
BASE COUNT 377 a 253 c 252 g 285 t
ORIGIN
polya_signal 1097..1102

```

Query Match 1.6%; Score 19; DB 59; Length 1167;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 tgagtaatggtgtaaaata 204  
 ||||||||||||||||  
 Db 191 TGAGTAATGCTGTAATAAATA 209

```

RESULT 4
LOCUS A63886 388 bp DNA
DEFINITION Sequence 11 from Patent WO9723633.
ACCESSION A63886
VERSION A63886.1 GI:3717432
KEYWORDS
SOURCE Phaffia rhodozyma.
ORGANISM Phaffia rhodozyma
REFERENCE 1 (bases 1 to 388)
AUTHORS Verdoes, J. C. and Wery, J.
TITLE IMPROVED METHODS FOR TRANSFORMING PHAFFIA STRAINS, TRANSFORMED
            PHAFFIA STRAINS SO OBTAINED AND RECOMBINANT DNA IN SAID METHODS
            Patent: WO 9723633-A 11 03-JUL-1997;
            GIST BROCADES BV (NL)
            Other publication AU 1308797, 19970717.
COMMENT Other publication AU 1308797, 19970717.
FEATURES
            location/Qualifiers
                1..388
                /organism="Phaffia rhodozyma"
                /db_xref="taxon:103478"
                1..385
                promoter
                TATA_signal 249..263
                misc_signal 287..302
                misc_RNA 386..388
                /function="START OF CDS"
BASE COUNT 97 a 121 c 77 g 92 t 1 others
ORIGIN

```

Query Match 1.5%; Score 18; DB 9; Length 388;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 400 aaagatctctctctca 417  
 ||||||||||||||||  
 Db 312 AAAGATCTCTCTCTTCA 329

```

RESULT 5
LOCUS G14342/c 400 bp DNA
DEFINITION SHGC-8810 Human Homo sapiens STS genomic clone pg-4539, sequence
            tagged site.
ACCESSION G14342
VERSION G14342.1 GI:1130081
KEYWORDS STS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 400)
AUTHORS Olivier, M. and Cox, D. R.
TITLE Unpublished, Olivier, M., Cox, D. R. (2000)
JOURNAL Unpublished (2000)
COMMENT

```

Contact: Michael Olivier, David R. Cox  
 Stanford Human Genome Center  
 Stanford University School of Medicine  
 4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA  
 Tel: (650) 320-5800  
 Fax: (650) 320-5801  
 Email: olivier@shgc.stanford.edu  
 Primer A: AGCACTTGACATCTGTGCA  
 Primer B: TCAGACTGGATGATGCTTGA  
 STS size: 166  
 PCR Profile:

Initial incubation: 94 degrees C for 90 seconds  
 Denaturation: 94 degrees C for 15 seconds  
 Annealing: 62 degrees C for 23 seconds

9	18	1.5	908	54	CNS07DTH	AL440907 T7 end of
10	18	1.5	1015	53	CNS07BDY	AL437756 T3 end of
11	18	1.5	1019	12	AF115672	AF115672 Semperiv
12	18	1.5	1066	53	CNS07GCT	AL439011 T3 end of
13	17	1.4	28	45	E08984	E08984 PCR primer.
14	17	1.4	177	92	HS36DAR	Z55360 H.sapiens C
15	17	1.4	179	92	HS36DAR	Z55359 H.sapiens C
16	17	1.4	198	54	HS119XB4	Z66555 H.sapiens D
17	17	1.4	217	54	G55253	G55253 SHGC-100521
18	17	1.4	308	7	AF055505	AF055505 OrnlNorth
19	17	1.4	313	54	G09193	G09193 human STS C
20	17	1.4	339	54	G24945	G24945 human STS E
21	17	1.4	400	54	G16415	G16415 SHGC-11640
22	17	1.4	400	54	G16870	G16870 human STS S
23	17	1.4	400	54	G34430	G34430 human STS S
24	17	1.4	403	10	AX079453	AX079453 Sequence
25	17	1.4	438	53	BT095713	BT095713 Bos taurus
26	17	1.4	466	53	CNS060UJ8	AL415914 T7 end of
27	17	1.4	480	54	G52169	AL415914 T7 end of
28	17	1.4	504	54	G32208	G52169 SHGC-79375
29	17	1.4	508	15	URE291974	G32208 STS 8.3588F
30	17	1.4	520	7	AF259758	AF259758 Ovis arie
31	17	1.4	532	7	AF101399	AF101399 Equus cab
32	17	1.4	563	7	BT095983	BT095983 Bos taurus
33	17	1.4	582	54	HSJ100C7	AL159559 STS from
34	17	1.4	621	94	AF239743	AF239743 Mus muscu
35	17	1.4	652	54	G52308	G52308 SHGC-79361
36	17	1.4	706	8	AF220549	AF220549 Paratichit
37	17	1.4	778	12	AF027686	AF027686 Onodrychi
38	17	1.4	850	53	CNS07132	AL424437 T7 end of
39	17	1.4	862	53	CNS06060	AL407686 T3 end of
40	17	1.4	880	91	AY008428S2	AY008430 Cercopit
41	17	1.4	883	91	AY008417S5	AY008432 Hylobates
42	17	1.4	912	94	AF224719	AF224719 Mus muscu
43	17	1.4	915	53	CNS06VFP	AL417433 T7 end of
44	17	1.4	923	53	CNS01EXD	AL411122 Anopheles
45	17	1.4	924	58	AF344647	AF344647 Papaya ri

## ALIGNMENTS

RESULT 1  
AX047405/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AX047405 609 bp DNA  
Sequence 21 from Patent W00068402.  
AX047405  
AX047405.1 GI:11876631

PAT 15-DEC-2000

bread wheat.  
Triticum aestivum  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticaceae;  
Triticum.  
1 (bases 1 to 609)  
van den Ende, W., van laere, A., de Roover, J. and Michiels, A.  
TITLE  
Manipulation of fructan catabolism in plants  
JOURNAL  
Patent: WO 0068402-A 21 16-NOV-2000;  
K.U. Leuven Research & Development (BE)  
LOCATION/Qualifiers

FEATURES  
source  
1..609  
/organism="Triticum aestivum"  
/db\_xref="taxon:4565"

BASE COUNT 159 a 121 c 164 g 165 t

ORIGIN

Query Match 1.6%; Score 20; DB 9; Length 609;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1148 caaagaatgatttgaagc 1167  
|||||

Db 119 CANAGATGATTGGAAGC 100

RESULT 2  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

G52287 654 bp DNA STS  
SHGC-79334 Human Homo sapiens STS genomic, sequence tagged site.  
G52287  
G52287.1 GI:5223464  
STS.  
human.  
Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 654)  
Olivier, M. and Cox, D.R. (2000)  
Unpublished (2000)  
Contact: Michael Olivier, David R. Cox  
Stanford Human Genome Center  
Stanford University School of Medicine  
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA  
Tel: (650) 320-5800  
Fax: (650) 320-5801  
Email: olivier@shgc.stanford.edu  
Primer A: CACCTGCACATCTAGGCTCT  
Primer B: GAGCTGACATCTTGAAGGCTA  
STS size: 310  
PCR profile:

Initial incubation: 95 degrees C for 10 minutes  
Denaturation: 94 degrees C for 30 seconds  
Annealing: 60 degrees C for 30 seconds  
Polymerization: 72 degrees C for 23 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9700

Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
AmpliTaq Gold Polymerase: 0.07 units/uL  
Total Vol: 5 uL

Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 8.3

FEATURES  
source  
1..654  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="4"  
/clone\_lib="Human"

STS  
primer\_bind  
primer\_bind  
BASE COUNT 196 a 146 c 103 g 209 t

ORIGIN

Query Match 1.6%; Score 19; DB 54; Length 654;  
Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 220 ctcaaacactgttaagact 238  
Db 624 CTCAAACACTGTTAAGACT 642  
|||||

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2001, 10:33:48 ; Search time 2170.02 Seconds  
(without alignments)  
8660.434 Million cell updates/sec

Title: US-09-389-000-1\_COPY\_735\_1949  
Perfect score: 1215  
Sequence: 1 ggtctcttgaagtgccctt.....ttgcacgtgttgcctctg 1215

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1344157 segs, 7733874588 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105172

Minimum DB seq length: 0

Maximum DB seq length: 1214

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_da1:\*  
2: gb\_da2:\*  
3: gb\_da3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_da1:\*  
17: em\_da2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rtd:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_inv3:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_rod:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_vl:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_vl1:\*  
59: gb\_vl2:\*  
60: gb\_htg1:\*  
61: gb\_htg2:\*  
62: gb\_htg3:\*  
63: gb\_htg4:\*  
64: gb\_htg5:\*  
65: gb\_htg6:\*  
66: gb\_htg7:\*  
67: gb\_htg8:\*  
68: gb\_htg9:\*  
69: gb\_htg10:\*  
70: gb\_htg11:\*  
71: gb\_htg12:\*  
72: gb\_htg13:\*  
73: gb\_htg14:\*  
74: gb\_htg15:\*  
75: gb\_htg16:\*  
76: gb\_htg17:\*  
77: gb\_htg18:\*  
78: gb\_htg19:\*  
79: gb\_htg20:\*  
80: gb\_htg21:\*  
81: gb\_htg22:\*  
82: gb\_htg23:\*  
83: gb\_htg24:\*  
84: gb\_htg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_r01:\*  
95: gb\_r02:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_da3:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	20	1.6	609	9	AX047405	AX047405 Sequence
2	19	1.6	654	54	G52287	G52287 SHGC-79334
3	19	1.6	1167	59	XCU070906	U70906 Xestia C-nl
4	18	1.5	388	9	A63886	A63886 Sequence 11
5	18	1.5	400	54	G14342	G14342 SHGC-8810 H
6	18	1.5	429	94	AB04300510	AB043018 Mus muscu
7	18	1.5	442	54	G51525	G51525 SHGC-80856
8	18	1.5	882	53	CNS06WRW	AL418818 T7 end of

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;  
 OC Moronidae; Morone.  
 OX NCBI\_TaxID=34816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=RETINA;  
 RX MEDLINE=96012164; PubMed=7553903;  
 RA King-Smith C., Bost-Usinger L., Burnside B.;  
 RT "Expression of kinesin heavy chain isoforms in retinal pigment  
 epithelial cells.";  
 RL Cell Motil. Cytoskeleton 31:66-81(1995).  
 DR EMBL: U34653; AAA85772.1; -;  
 DR HSP; P33176; IBC2.  
 DR InterPro; IPR001752; -;  
 DR Pfam; PF00225; kinesin; 1.  
 DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 DR SMART; SM00129; KISC; 1.  
 KW ATP-binding; Coiled coil; Microtubules; Motor protein.  
 FT NON\_TER 1  
 FT NON\_TER 145 145  
 SQ SEQUENCE 145 AA; 16710 MW; 44A6091EAC23487B CRC64;

Query Match 71.4%; Score 5; DB 13; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PYVKG 5  
 Db 74 PYVKG 78

Search completed: September 15, 2001, 12:50:18  
 Job time: 278 sec

OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.  
 OX NCBI\_TaxID=5762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Burger G., Lang B.F., Nerad T.A., Gray M.W.;  
 RT "The mitochondrial genome of the supposedly primitive protist,  
 RT Naegleria gruberi.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF288092; AAG1782.1; -;  
 KW Mitochondrion.  
 SQ SEQUENCE 128 AA; 14393 MW; D0E4B6E8AD941400 CRC64;

Query Match 71.4%; Score 5; DB 8; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
 Db 108 VKGRK 112  
 |||||

RESULT 13  
 Q9VXR0 PRELIMINARY; PRT; 131 AA.  
 ID Q9VXR0;  
 AC Q9VXR0;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE CG8198 PROTEIN.  
 GN CG8198.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
 RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M.C., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003500; AAF48498.1; -;  
 DR FlyBase; FBgn0030677; CG8198.  
 DR InterPro; IPR000361; -;  
 DR Pfam; PF01521; HesB-like; 1.  
 DR PROSITE; PS01152; HESB; 1.  
 SQ SEQUENCE 131 AA; 14211 MW; BCF760E25FAF7426 CRC64;

Query Match 71.4%; Score 5; DB 5; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
 Db 14 VKGRK 18  
 |||||

RESULT 14  
 Q9HLY2 PRELIMINARY; PRT; 142 AA.  
 ID Q9HLY2;  
 AC Q9HLY2;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE PROBABLE RIBOSOMAL PROTEIN S12.  
 GN TA0091.  
 OS Thermoplasma acidophilum.  
 OC Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaaceae;  
 OC Thermoplasma.  
 OX NCBI\_TaxID=2303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 1728;  
 RX MEDLINE=20479972; PubMed=11029001;  
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,  
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;  
 RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma  
 RT acidophilum".  
 RL Nature 407:508-513(2000).  
 DR EMBL: AL445063; CAC11239.1; -;  
 DR InterPro; IPR000230; -;  
 DR Pfam; PF00164; Ribosomal\_S12; 1.  
 DR PROSITE; PS00055; RIBOSOMAL\_S12; 1.  
 KW Ribosomal protein.  
 SQ SEQUENCE 142 AA; 15751 MW; BC90C9EEDA2C52BA CRC64;

Query Match 71.4%; Score 5; DB 1; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
 Db 133 VKGRK 137  
 |||||

RESULT 15  
 Q91110 PRELIMINARY; PRT; 145 AA.  
 ID Q91110;  
 AC Q91110;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE KINESIN HEAVY CHAIN FKFI1 (FRAGMENT).  
 GN FKFI1.  
 OS Morone saxatilis (Striped bass).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;



Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVKG 6

Db 26 YVKG 30

RESULT 9

Q9HXV2 ID Q9HXV2 PRELIMINARY; PRT; 114 AA.

AC Q9HXV2;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

DE HYPOTHETICAL PROTEIN PA3688.

GN PA3688.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI\_TaxID=287;

RP [1]

RN SEQUENCE FROM N.A.

RC STRAIN=PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,

RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Goutry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,

RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen."

RL Nature 406:959-964(2000).

DR EMBL; AE004788; AAG07076.1; -.

KW Hypothetical protein.

SQ SEQUENCE 114 AA; 12918 MW; E3E4C7EFE983E268 CRC64;

Query Match

Best Local Similarity 71.4%; Score 5; DB 2; Length 114;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVKG 5

Db 67 PVKG 71

RESULT 10

P78570 ID P78570 PRELIMINARY; PRT; 114 AA.

AC P78570;

DT 01-MAY-1997 (TrEMBLrel. 03, Created)

DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

DE DNA-DIRECTED RNA POLYMERASE I (EC 2.7.6) (RNA POLYMERASE II) (RNA

DE POLYMERASE III) (RNA NUCLEOTIDYLTRANSFERASE (DNA-DIRECTED))

DE (FRAGMENT).

GN RPAB.

OS Agaricus bisporus (Common mushroom).

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Agaricales;

OC Agaricaceae; Agaricus.

OX NCBI\_TaxID=5341;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HORST U1.

RX MEDLINE=97111990; PubMed=8953726;

RA Sonnenberg A.S.M., De Groot P.W.J., Schaap P.J., Baars J.J.P.,

RA Visser J., van Griensven L.J.B.D.;

RT "Isolation of expressed sequence tags of Agaricus bisporus and their

RT assignment to chromosomes."

RL Appl. Environ. Microbiol. 62:4542-4547(1996).

CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE -> N PYROPHOSPHATE +

CC RNA(N) (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

DR EMBL; X94765; CRA64391.1; -.

DR InterPro; IPR001572; -.

DR Pfam; PF00562; RNA\_pol\_B; 1.

KW DNA-directed RNA polymerase; Nucleotidyltransferase; Transcription;

KW Transferase.

FT NON\_TER 1 1

FT NON\_TER 114 114

SQ SEQUENCE 114 AA; 12903 MW; 656C52B0AE6C2908 CRC64;

Query Match

Best Local Similarity 71.4%; Score 5; DB 3; Length 114;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7

Db 82 VKGRK 86

RESULT 11

Q9SXY4

ID Q9SXY4 PRELIMINARY; PRT; 124 AA.

AC Q9SXY4;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

DE CLASS III CHITINASE HOMOLOGUE (OSCHIB3H-G) (FRAGMENT).

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;

OC Oryza.

OX NCBI\_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RA Itoh Y., Sasaki T.;

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=98162724; PubMed=9501993;

RA Nagasaki H., Yamamoto K., Shomura A., Koga-Ban Y., Takasuga A.,

RA Yano M., Minobe Y., Sasaki T.;

RT "Rice class III chitinase homologues isolated by random cloning of

RT rice cDNAs."

RL DNA Res. 4:379-385(1997).

DR EMBL; AB027426; BAA7779.1; -.

DR HSSP; P23472; 2HVM.

DR InterPro; IPR001579; -.

DR Pfam; PF00192; chitinase\_2; 2.

FT NON\_TER 124 124

SQ SEQUENCE 124 AA; 13193 MW; 6751E605EC8A4D6E CRC64;

Query Match

Best Local Similarity 71.4%; Score 5; DB 10; Length 124;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVKG 6

Db 63 YVKG 67

RESULT 12

Q9G8S2

ID Q9G8S2 PRELIMINARY; PRT; 128 AA.

AC Q9G8S2;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

DE RIBOSOMAL PROTEIN S12.

GN RPS12.

OS Naegleria gruberi.

OG Mitochondrion.

```
QY 1 PVVGR 6
Db 460 PVVGR 465

RESULT 5
O28788 PRELIMINARY; PRT; 53 AA.
AC O28788;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 5.6 KDA PROTEIN.
GN AF1484.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98049343; PubMed=9389475;
RX Klenk H.-P., Clayton R.A., Gwinn M., Hickey E.K., Peterson J.D.,
RA Ketchum K.A., Dodson R.J., Tomb J.-F., White O., Nelson K.E.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AF001000; AAB89773.1; -.
DR TIGR; AF1484; -.
KW Hypothetical protein.
SQ SEQUENCE 53 AA; 5609 MW; 729877365F2639D0 CRC64;

Query Match 71.4%; Score 5; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVGR 6
Db 18 YVGR 22

RESULT 6
Q63212 PRELIMINARY; PRT; 59 AA.
AC Q63212;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GAP-43 PROTEIN.
GN GAP43.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=92185515; PubMed=1532026;
RA Nedivi E., Basi G.S., Akey I.V., Skene J.H.P.;
RT "A neural-specific GAP-43 core promoter located between unusual DNA
RT elements that interact to regulate its activity.";
RL J. Neurosci. 12:691-704(1992).
```

---

```
DR EMBL; M88356; AAA41191.1; -.
DR InterPro; IPR001422; -.
DR PROSITE; PS00412; NEUROMODULIN_1; 1.
SQ SEQUENCE 59 AA; 7024 MW; 154511D1856977A1 CRC64;

Query Match 71.4%; Score 5; DB 11; Length 59;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 7
Db 32 VKGR 36

RESULT 7
Q9N7S8 PRELIMINARY; PRT; 71 AA.
AC Q9N7S8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 8.1 KDA PROTEIN (FRAGMENT).
GN LM28.104.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL390935; CAC00859.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 8066 MW; A4D6A4D3E51520DF CRC64;

Query Match 71.4%; Score 5; DB 5; Length 71;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGR 7
Db 7 VKGR 11

RESULT 8
Q9N7F8 PRELIMINARY; PRT; 86 AA.
AC Q9N7F8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE POSSIBLE GCD14 PROTEIN (FRAGMENT).
GN LM28.263.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL390935; CAC00979.1; -.
FT NON_TER 1
FT NON_TER 86
SQ SEQUENCE 86 AA; 9482 MW; F95324DBBDF5191A CRC64;

Query Match 71.4%; Score 5; DB 5; Length 86;
Best Local Similarity 100.0%; Pred. No. 78;
```

GN Y37A1B.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Mcmurray A.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smaldon N., Smith A., Sonnenhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; AL023835; CAA19490.1; --  
 DR InterPro; IPR000345; --  
 DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
 SQ SEQUENCE 471 AA; 52424 MW; C4E95E16614E1C50 CRC64;

Query Match 85.7%; Score 6; DB 5; Length 471;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YVKGK 7  
 DB 376 YVKGK 391  
 ID 044507 PRELIMINARY; PRT; 487 AA.  
 AC 044507;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE F42G8.8 PROTEIN.  
 GN F42G8.8.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smaldon N., Smith A., Sonnenhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; AL023835; CAA19490.1; --  
 DR InterPro; IPR000345; --  
 DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
 SQ SEQUENCE 471 AA; 52424 MW; C4E95E16614E1C50 CRC64;

RA Gattung S., Holmes A.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: A PHOSPHOPROTEIN + H(2)O = A PROTEIN +  
 CC ORTHOPHOSPHATE (THIS ENZYME IS SERINE/THREONINE SPECIFIC) (BY  
 CC SIMILARITY).  
 CC -!- SIMILARITY: TO SERINE/THREONINE SPECIFIC PROTEIN PHOSPHATASE (OR  
 CC PPP) FAMILY.  
 DR EMBL; AF038618; AAB92072.1; --  
 DR HSSP; P08129; 1FUM.  
 DR InterPro; IPR000934; --  
 DR Pfam; PF00149; Stphosphatase; 1.  
 DR PRINTS; PR00114; STPHPHATASE.  
 DR PROSITE; PS00125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR SMART; SM00156; PP2Ac; 1.  
 KW Hydrolase; Iron; Manganese.  
 SQ SEQUENCE 487 AA; 54922 MW; 3D94365514393C40 CRC64;

Query Match 85.7%; Score 6; DB 5; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YVKGK 7  
 DB 445 YVKGK 450  
 ID 09RIU7 PRELIMINARY; PRT; 772 AA.  
 AC 09RIU7;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
 DE PUTATIVE SECRETED PROTEIN.  
 GN SCML1.14C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Oliver K., Harris D.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb streptomycetes coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL; AL133278; CAB61925.1; --  
 DR InterPro; IPR001899; --  
 DR PROSITE; PS00343; GRAM\_POS\_ANCHORING; UNKNOWN\_1.  
 SQ SEQUENCE 772 AA; 84256 MW; 6B5DAF10237CF460 CRC64;

Query Match 85.7%; Score 6; DB 2; Length 772;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:50:13 ; Search time 69.13 seconds  
(without alignments)  
13.397 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_163\_169  
Perfect score: 7  
Sequence: 1 PVYGRK 7

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : SPTREMBL\_16:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_plage:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_unclassified:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	100.0	425	4 Q9NX45	Q9NX45 homo sapien
2	6	85.7	471	5 Q9XXF9	Q9XXF9 caenorhabdi
3	6	85.7	487	5 O44507	O44507 caenorhabdi
4	6	85.7	772	2 Q9RI07	Q9RI07 streptomyce
5	5	71.4	53	1 O28788	O28788 archaeoglob
6	5	71.4	59	11 Q63212	Q63212 rattus norv
7	5	71.4	71	5 Q9N7S8	Q9N7S8 leishmania
8	5	71.4	86	5 Q9N7F8	Q9N7F8 leishmania
9	5	71.4	114	2 Q9HXV2	Q9HXV2 pseudomonas
10	5	71.4	114	3 P78570	P78570 agaricus bi
11	5	71.4	124	10 Q9SXY4	Q9SXY4 oryza sativ
12	5	71.4	128	8 Q9G8S2	Q9G8S2 naegleria g
13	5	71.4	131	5 Q9VXR0	Q9VXR0 drosophila
14	5	71.4	142	1 Q9HLY2	Q9HLY2 thermoplas
15	5	71.4	145	13 Q91110	Q91110 morone saxa
16	5	71.4	147	1 Q9V110	Q9V110 pyrococcus
17	5	71.4	148	10 Q9O801	Q9O801 arabidopsis
18	5	71.4	157	8 Q9MTQ3	Q9MTQ3 amphidinium
19	5	71.4	162	10 Q22325	Q22325 musa acumin

20	5	71.4	162	14	Q02476	Q02476 poa semilat
21	5	71.4	163	2	Q9RF18	Q9RF18 amoa anoxic
22	5	71.4	176	6	Q9N2G6	Q9N2G6 sus scrofa
23	5	71.4	178	6	Q9TVB3	Q9TVB3 bos taurus
24	5	71.4	186	2	Q47076	Q47076 escherichia
25	5	71.4	187	10	O48515	O48515 ipomoea bat
26	5	71.4	187	10	O48516	O48516 ipomoea bat
27	5	71.4	190	5	Q9GYC2	Q9GYC2 leishmania
28	5	71.4	194	2	O50912	O50912 borrelia bu
29	5	71.4	197	2	O32146	O32146 bacillus su
30	5	71.4	221	2	Q9F5K9	Q9F5K9 streptomyce
31	5	71.4	225	2	Q9KJZ8	Q9KJZ8 pseudomonas
32	5	71.4	235	8	Q33368	Q33368 chlamydomon
33	5	71.4	235	10	O98224	O98224 chlamydomon
34	5	71.4	243	10	Q9FEH5	Q9FEH5 chlamydomon
35	5	71.4	243	11	Q9UKA4	Q9UKA4 mus musculu
36	5	71.4	245	11	Q9UKA1	Q9UKA1 mus musculu
37	5	71.4	252	5	Q23327	Q23327 caenorhabdi
38	5	71.4	254	2	Q9KXX2	Q9KXX2 streptomyce
39	5	71.4	256	2	O52759	O52759 rhizobium l
40	5	71.4	262	2	Q9FAB6	Q9FAB6 streptomyce
41	5	71.4	267	10	Q9SSX7	Q9SSX7 nicotiana t
42	5	71.4	278	2	O44579	O44579 acetobacter
43	5	71.4	278	2	O44580	O44580 acetobacter
44	5	71.4	278	2	O48339	O48339 acetobacter
45	5	71.4	287	3	Q9USH0	Q9USH0 schizosacch

ALIGNMENTS

RESULT 1

Q9NX45 ID Q9NX45 PRELIMINARY; PRT; 425 AA.

AC Q9NX45;

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE CDNA FLJ20449 FIS, CLONE KAT05575.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T., Suzuki Y., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;

RT "NEDO human cDNA sequencing project."

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK000456; BAA91175.1; -.

DR InterPro; IPR001092; -.

DR SMART; SM00353; HLH; 1.

SQ SEQUENCE 425 AA; 46922 MW; 777E57231F9D6BDD CRC64;

Query Match 100.0%; Score 7; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVYGRK 7  
DB 229 PVYGRK 235

RESULT 2

Q9XXF9 ID Q9XXF9 PRELIMINARY; PRT; 471 AA.

AC Q9XXF9;

DT 01-NOV-1999 (Tremblrel. 12, Created)

DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)

DE Y37A1B.5 PROTEIN.



DR InterPro; IPR000498; -  
DR InterPro; IPR001145; -  
DR Pfam; PF00691; OmpA; 1.  
DR Pfam; PF01389; OmpA\_membrane; 1.  
DR PRINTS; PR01021; OMPADOMAIN.  
DR PROSITE; PS01068; OMPA; 1.  
KW Outer membrane; Transmembrane; Porin; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 353 OUTER MEMBRANE PROTEIN P5.  
FT DISULFID 326 338 BY SIMILARITY.  
FT DOMAIN 272 316 OMPA-LIKE.  
SQ SEQUENCE 353 AA; 37594 MW; E59A659E7860D0F7 CRC64;

Query Match 71.4%; Score 5; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
|||||  
Db 329 VKGRK 333

RESULT 14  
OM53\_HAEIN STANDARD; PRT; 359 AA.  
AC P45996;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE OUTER MEMBRANE PROTEIN P5 PRECURSOR (OMP P5) (FIMBRIN).  
GN OMPA OR OMP5.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC STRAIN=NTHT 1128;  
RX MEDLINE=9422575; PubMed=7909539;  
RA Sirakova T., Kolattukudy P.E., Murwin D., Billy J., Leake E.,  
RA Lim D., Demaria T., Bakalez L.;  
RT "Role of fimbriae expressed by nontypeable Haemophilus influenzae in  
pathogenesis of and protection against otitis media and relatedness  
of the fimbria subunit to outer membrane protein A.";  
RL Infect. Immun. 62:2002-2020(1994).  
CC -!- FUNCTION: ACTS AS A FIMBRIAE SUBUNIT.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.  
CC -!- SIMILARITY: BELONGS TO THE OMPA FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; L08448; AAA24959.1; -  
DR InterPro; IPR000498; -  
DR InterPro; IPR001145; -  
DR Pfam; PF00691; OmpA; 1.  
DR Pfam; PF01389; OmpA\_membrane; 1.  
DR PRINTS; PR01021; OMPADOMAIN.  
DR PROSITE; PS01068; OMPA; FALSE\_NEG.  
KW Outer membrane; Transmembrane; Porin; Signal; Fimbria.  
FT SIGNAL 1 21  
FT CHAIN 22 359 OUTER MEMBRANE PROTEIN P5.  
FT DISULFID 332 344 BY SIMILARITY.  
FT DOMAIN 278 322 OMPA-LIKE.  
SQ SEQUENCE 359 AA; 38340 MW; 576B1C59B4818C37 CRC64;

Query Match 71.4%; Score 5; DB 1; Length 359;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
|||||  
Db 335 VKGRK 339

RESULT 15  
YBWL\_CAEEL STANDARD; PRT; 364 AA.  
AC Q18347;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE HYPOTHETICAL 41.5 KDA PROTEIN C31H2.4 IN CHROMOSOME X.  
GN C31H2.4.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Geisel C., Gattung S.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE 4HPPD FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U41748; AAA83338.1; -  
DR WormPep; C31H2.4; CE04132.  
DR InterPro; IPR002887; -  
DR Pfam; PF01626; 4HPPD\_C; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 364 AA; 41456 MW; 599310F028F28E18 CRC64;

Query Match 71.4%; Score 5; DB 1; Length 364;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
|||||  
Db 225 VKGRK 229

Search completed: September 15, 2001, 12:50:43  
Job time: 298 sec

RT degrading enzymes from the thermophile 'Caldocellum  
RT saccharolyticum';  
RL Appl. Environ. Microbiol. 56:1017-1024(1990).  
RN [2].  
RP SEQUENCE FROM N.A.  
RA Te'o V.S. Jr., Gibbs M.D., Saul D.J., Bergquist P.L.;  
RT "A cluster of genes involved in xylan degradation cloned from the  
RT extreme thermophile Caldicellulosiruptor saccharolyticus";  
RL Appl. Environ. Microbiol. 0:0-0(1997).  
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC  
CC LINKAGES IN XYLANS.  
CC -!- PATHWAY: XILAN DEGRADATION.  
CC -!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL  
CC HYDROLASES).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M34459; AAA23059.1; -;  
DR EMBL; AF005383; AAB97374.1; -;  
DR PIR; A37202; A37202.  
DR HSSP; P56588; 1BG4.  
DR InterPro: IPR001000; -;  
DR Pfam; PF00331; Glyco\_hydro\_10; 1.  
DR PRINTS; PR00134; GLHYDRLASE10.  
DR PROSITE; PS00591; GLYCOSYL\_HYDROL\_F10; 1.  
KW Xylan degradation; Hydrolase; Glycosidase; Signal.  
FT SIGNAL 1 33  
FT CHAIN 34 342 ENDO-1,4-BETA-XYLANASE A.  
FT ACT\_SITE 144 144 PROTON DONOR (BY SIMILARITY).  
FT ACT\_SITE 252 252 NUCLEOPHILE (BY SIMILARITY).  
FT ACT\_SITE 252 252 NUCLEOPHILE (BY SIMILARITY).  
SQ SEQUENCE 342 AA; 40455 MW; C5380F2AB0CC0271 CRC64;  
  
Query Match 71.4%; Score 5; DB 1; Length 342;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 3 VKGRK 7  
Db 316 VKGRK 320  
|||||  
  
RESULT 12  
OM51\_HAEIN STANDARD; PRT; 353 AA.  
ID OM51\_HAEIN STANDARD; PRT; 353 AA.  
AC P43840;  
DT 01-NOV-1995 (Rel. 32; Created)  
DT 01-NOV-1995 (Rel. 32; Last sequence update)  
DT 01-NOV-1995 (Rel. 32; Last annotation update)  
DE OUTER MEMBRANE PROTEIN P5 PRECURSOR (OMP P5).  
GN OMPA OR OMP5 OR H1164.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OC NCBI\_TaxID=727;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=RD / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus  
RT influenzae Rd.";  
RL Science 269:496-512(1995).  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.  
CC -!- SIMILARITY: BELONGS TO THE OMPA FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U32796; AAC22819.1; -;  
DR TIGR; H11164; -;  
DR InterPro; IPR000498; -;  
DR InterPro; IPR001145; -;  
DR Pfam; PF00691; OmpA; 1.  
DR Pfam; PF01389; OmpA\_membrane; 1.  
DR PRINTS; PR01021; OMPADOMAIN.  
DR PROSITE; PS01066; OMPA; 1.  
KW Outer membrane; Transmembrane; Porin; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 353 BY SIMILARITY.  
FT DISULFID 326 338 OUTER MEMBRANE PROTEIN P5.  
FT DOMAIN 272 316 BY SIMILARITY.  
FT DOMAIN 316 316 OMPA-LIKE.  
SQ SEQUENCE 353 AA; 37743 MW; 64ACB3E7BFF96B39 CRC64;  
  
Query Match 71.4%; Score 5; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 3 VKGRK 7  
Db 329 VKGRK 333  
|||||  
  
RESULT 13  
OM52\_HAEIN STANDARD; PRT; 353 AA.  
ID OM52\_HAEIN STANDARD; PRT; 353 AA.  
AC P38368;  
DT 01-OCT-1994 (Rel. 30; Created)  
DT 01-OCT-1994 (Rel. 30; Last sequence update)  
DT 01-NOV-1995 (Rel. 32; Last annotation update)  
DE OUTER MEMBRANE PROTEIN P5 PRECURSOR (OMP P5).  
GN OMPA OR OMP5.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OC NCBI\_TaxID=727;  
OX [1]  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 22-34.  
RP STRAIN=1613 / SEROTYPE B;  
RX MEDLINE=93366472; PubMed=8359929;  
RA Munson R.S. Jr., Grass S., West R.;  
RT "Molecular cloning and sequence of the gene for outer membrane  
RT protein P5 of Haemophilus influenzae";  
RL Infect. Immun. 61:4017-4020(1993).  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.  
CC -!- SIMILARITY: BELONGS TO THE OMPA FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L20309; AAA03346.1; -;

DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 DR PROSITE; PS00067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KW Motor protein; Microtubules; ATP-binding; Coiled coil; 3D-structure.  
 FT INIT\_MET 0  
 FT DOMAIN 1 >338 MECHANOCHEMICAL (MOTOR) (BY SIMILARITY).  
 FT DOMAIN 173 >238 MICROTUBULE-BINDING.  
 FT NP\_BIND 85 92 ATP (BY SIMILARITY).  
 FT NON\_TER 238 238  
 SQ SEQUENCE 238 AA; 27034 MW; 7B254F4894E3BC30 CRC64;

Query Match 71.4%; Score 5; DB 1; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVYKG 5

Db 163 PVYKG 167

# RESULT 9

ID VATE\_METTL STANDARD; PRT; 253 AA.  
 AC P20022;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE V-TYPE ATP SYNTHASE BETA CHAIN (EC 3.6.1.34) (V-TYPE ATPASE SUBUNIT B) (FRAGMENT).  
 DE B (FRAGMENT).  
 GN ATPB.  
 OS Methanococcus thermolithotrophicus.  
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
 OC Methanococcus.  
 OX NCBI\_TaxID=2186;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89325654; PubMed=2526753;  
 RA Bernasconi P., Rausch T., Gogarten J.P., Taiz L.;  
 RT "The H<sup>+</sup> ATPase regulatory subunit of Methanococcus  
 thermolithotrophicus: amplification of an 800 bp fragment by  
 polymerase chain reaction.";  
 RL FEBS Lett. 251:132-136(1989).  
 CC -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON  
 GRADIENT ACROSS THE MEMBRANE. THE ARCHAEAL BETA CHAIN IS A  
 REGULATORY SUBUNIT.  
 CC -!- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.  
 DR PIR; S05029; S05029.  
 DR InterPro; IPR000194; -.  
 DR Pfam; PF00006; ATP-synt\_ab; 1.  
 DR PROSITE; PS00152; ATPASE\_ALPHA\_BETA; PARTIAL.  
 KW Hydrolase; ATP synthesis; Hydrogen ion transport.  
 FT NON\_TER 1 1  
 FT NON\_TER 253 253  
 SQ SEQUENCE 253 AA; 27669 MW; F3EF9582D0851AD1 CRC64;

Query Match 71.4%; Score 5; DB 1; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7

Db 228 VKGRK 232

# RESULT 10

ID NECD\_MOUSE STANDARD; PRT; 325 AA.  
 AC P25233; Q61951;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE NECDIN.

GN NDN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91298962; PubMed=2069569;  
 RA Maruyama K., Usami M., Aizawa T., Yoshikawa K.;  
 RT "A novel brain-specific mRNA encoding nuclear protein (necdin)  
 expressed in neurally differentiated embryonal carcinoma cells.";  
 RL Biochem. Biophys. Res. Commun. 178:291-296(1991).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SV.  
 RX MEDLINE=96132826; PubMed=8557705;  
 RA Uetsuki T., Takagi K., Sugiura H., Yoshikawa K.;  
 RT "Structure and expression of the mouse necdin gene. Identification of  
 a postmitotic neuron-restrictive core promoter.";  
 RL J. Biol. Chem. 271:918-924(1996).  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURALLY DIFFERENTIATED EMBRYONAL  
 CARCINOMA CELLS AND IN ADULT BRAIN.  
 CC -!- SIMILARITY: TO HUMAN MELANOMA-ASSOCIATED ANTIGENS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; M80840; AAA39805.1; -.  
 CC EMBL; D76440; BAA11183.1; -.  
 CC PIR; JN0148; JN0148.  
 CC MGD; MGI:97290; Ndn.  
 CC InterPro; IPR002190; -.  
 CC Pfam; PF01454; MAGE; 1.  
 KW Nuclear protein; Brain.  
 FT VARIANT 17 17 V -> A.  
 FT SEQUENCE 325 AA; 36831 MW; 777385B0B75E443F CRC64;  
 SQ  
 Query Match 71.4%; Score 5; DB 1; Length 325;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 YVYGR 6  
 Db 203 YVYGR 207  
 RESULT 11  
 ID XYNA\_CALSA STANDARD; PRT; 342 AA.  
 AC P23556;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE ENDO-1,4-BETA-XYLANASE A PRECURSOR (EC 3.2.1.8) (XYLANASE A)  
 DE (1,4-BETA-D-XYLAN XYLANOHYDROLASE A).  
 GN XYNA.  
 OS Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Thermoanaerobacter group; Caldicellulosiruptor.  
 OX NCBI\_TaxID=44001;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90253140; PubMed=2111111;  
 RA Luethi E., Love D.R., McAnulty J., Wallace C., Caughey P.A.,  
 RA Saul D., Bergquist P.L.;  
 RT "Cloning, sequence analysis, and expression of genes encoding xylan-



DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DE 01-OCT-2000 (Rel. 40, Last annotation update)  
DE INTERLEUKIN-7 PRECURSOR (IL-7).  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89098903; PubMed=2643102;  
RA Goodwin R.G., Lupton S., Schmierer A., Hjerrild K.J., Jerzy R.,  
Clevenger W., Gillis S., Cosman D., Namen A.E.;  
RT "Human interleukin 7: molecular cloning and growth factor activity on  
human and murine B-lineage cells."  
RL Proc. Natl. Acad. Sci. U.S.A. 86:302-306(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90229755; PubMed=2329282;  
RA Lupton S.D., Gimpel S., Jerzy R., Brunton L.L., Hjerrild K.A.,  
Cosman D., Goodwin R.G.;  
RT "Characterization of the human and murine IL-7 genes."  
RL J. Immunol. 144:3592-3601(1990).  
RN [3]  
RP DISULFIDE BONDS, AND MASS SPECTROMETRY.  
RX MEDLINE=98070497; PubMed=9407080;  
RA Cosenza L., Sweeney E., Murphy J.R.;  
RT "Disulfide bond assignment in human interleukin-7 by matrix-assisted  
laser desorption/ionization mass spectroscopy and site-directed  
cysteine to serine mutational analysis."  
RL J. Biol. Chem. 272:32995-33000(1997).  
RN [4]  
RP 3D-STRUCTURE MODELING.  
RX MEDLINE=97015912; PubMed=862549;  
RA Kroemer R.T., Doughty S.W., Robinson A.J., Richards W.G.;  
RT "Prediction of the three-dimensional structure of human interleukin-7  
by homology modelling."  
RL protein Eng. 9:493-498(1996).  
RN [5]  
RP 3D-STRUCTURE MODELING.  
RX MEDLINE=20306974; PubMed=10850801;  
RA Cosenza L., Rosenbach A., White J.V., Murphy J.R., Smith T.F.;  
RT "Comparative model building of interleukin-7 using interleukin-4 as a  
template: a structural hypothesis that displays atypical surface  
chemistry in helix D important for receptor activation."  
RL Protein Sci. 9:916-926(2000).  
CC -!- FUNCTION: HEMATOPOIETIC GROWTH FACTOR CAPABLE OF STIMULATING THE  
PROLIFERATION OF LYMPHOID PROGENITORS. IT IS IMPORTANT FOR  
PROLIFERATION DURING CERTAIN STAGES OF B-CELL MATURATION.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- SIMILARITY: BELONGS TO THE IL-7 / IL-9 FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; J04156; AAA59156.1; -  
DR EMBL; M29053; AAC63047.1; -  
DR EMBL; M29048; AAC63047.1; JOINED.  
DR EMBL; M29049; AAC63047.1; JOINED.  
DR EMBL; M29050; AAC63047.1; JOINED.  
DR EMBL; M29051; AAC63047.1; JOINED.  
DR EMBL; M29052; AAC63047.1; JOINED.  
DR PIR; A32223; A32223.  
DR PIR; B32223; B32223.  
DR PIR; C32223; C32223.  
DR PIR; A43527; A43527.  
DR PDB; 1IL7; 04-MAR-98.  
DR MIM; 146660; -

DR InterPro; IPR000226; -  
DR InterPro; IPR001181; -  
DR Pfam; PF01415; IL7; 1.  
DR PRINTS; PR00435; INTERLEUKIN7.  
DR PROSITE; PS00255; INTERLEUKIN\_7\_9; 1.  
KW Cytokine; Growth factor; Glycoprotein; Signal; 3D-structure.  
FT SIGNAL 1 25  
FT CHAIN 26 177 INTERLEUKIN-7.  
FT DISULFID 27 166  
FT DISULFID 59 154  
FT DISULFID 72 117  
FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 177 AA; 20186 MW; 8FC5243F9169617F CRC64;  
  
Query Match 71.48; Score 5; DB 1; Length 177;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 3 VKGRK 7  
Db 121 VKGRK 125  
| | | | |  
| | | | |  
RESULT 8  
ID KF5C\_RAT STANDARD; PRO; 238 AA.  
AC P56536;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DE 01-OCT-2000 (Rel. 40, Last annotation update)  
DE KINESIN HEAVY CHAIN ISOFORM 5C (KINESIN HEAVY CHAIN NEURON-SPECIFIC  
DE 2) (FRAGMENT).  
GN KIF5C OR NKHC2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RC TISSUE=Brain;  
RX MEDLINE=96070240; PubMed=9405049;  
RA Sack S., Mueller J., Marx A., Thormaehlen M., Mandelkow E.M.,  
Brady S.T., Mandelkow E.;  
RT "X-ray structure of motor and neck domains from rat brain kinesin."  
RL Biochemistry 36:16155-16165(1997).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS).  
RX MEDLINE=98088797; PubMed=9428521;  
RA Kozielski F., Sack S., Marx A., Thormaehlen M., Schonbrunn E., Biou V.,  
Thompson A., Mandelkow E.M., Mandelkow E.;  
RT "The crystal structure of dimeric kinesin and implications for  
microtubule-dependent motility."  
RL Cell 91:985-994(1997).  
CC -!- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING  
PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.  
CC -!- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT  
CHAINS.  
CC -!- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-  
TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF  
KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL  
ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN  
DIMERIZATION; AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH  
INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),  
VESICLES AND MEMBRANOUS ORGANELLES.  
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN  
SUBFAMILY.  
CC PDB; 2KIN; 08-APR-98.  
CC PDB; 3KIN; 14-OCT-98.  
DR InterPro; IPR001752; -  
DR Pfam; PF00225; kinesin; 1.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X60305; CAA42849.1; -;  
 DR EMBL; X67313; CAA47727.1; -;  
 DR PIR; S18713; S18713.  
 DR InterPro; IPR000230; -;  
 DR Pfam; PF00164; Ribosomal\_S12; 1.  
 DR PROSITE; PS00055; RIBOSOMAL\_S12; 1.  
 KW Ribosomal protein.  
 SQ SEQUENCE 147 AA; 16421 MW; C55DB32035C5818A CRC64;

Query Match 71.4%; Score 5; DB 1; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
 |||||  
 DB 138 VKGRK 142

RESULT 5  
 RS12\_PYRHO STANDARD; PRT; 150 AA.  
 AC O59229;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 30S RIBOSOMAL PROTEIN S12P.  
 GN RPS12P OR PH1542.  
 OS Pyrococcus horikoshii.  
 CC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 OX NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE=98344137; PubMed=9679194;  
 RA Kawarabayashi Y., Savada M., Horikawa H., Haikawa Y., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kushida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 CC -!- FUNCTION: PROTEIN S12 IS INVOLVED IN THE TRANSLATION INITIATION  
 CC STEP (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE S12P FAMILY OF RIBOSOMAL PROTEINS.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AP000006; BAA30632.1; -;  
 DR InterPro; IPR000230; -;  
 DR Pfam; PF00164; Ribosomal\_S12; 1.  
 DR PRINTS; PR01034; RIBOSOMALS12.  
 DR PROSITE; PS00055; RIBOSOMAL\_S12; FALSE\_NEG.  
 KW Ribosomal protein.  
 SQ SEQUENCE 150 AA; 16800 MW; 9432F6B19D8D9F34 CRC64;

Query Match 71.4%; Score 5; DB 1; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7  
 |||||  
 DB 141 VKGRK 145

RESULT 6  
 BVIE\_BETVE STANDARD; PRT; 159 AA.  
 AC P43178;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE MAJOR POLLEN ALLERGEN BET V 1-E (BET V I-E).  
 GN BETVIE.  
 OS Betula verrucosa (White birch) (Betula pendula).  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 CC Fagales; Betulaceae; Betula.  
 OX NCBI\_TaxID=3505;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Pollen;  
 RX MEDLINE=95155322; PubMed=7852325;  
 RA Swoboda I., Jilek A., Ferreira F., Engel E., Hoffman-Sommergruber K.,  
 RA Scheiner O., Kraft D., Breiteneder H., Pittenauer E., Schmid E.,  
 RA Vicente O., Heberle-Bors E., Ahorn H., Breitenbach M.;  
 RT "Isoforms of Bet v 1, the major birch pollen allergen, analyzed by  
 RT liquid chromatography, mass spectrometry, and cDNA cloning.";  
 RL J. Biol. Chem. 270:2607-2613(1995).  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -!- DISEASE: MAIN CAUSE OF TYPE I ALLERGIC REACTIONS IN EUROPE, NORTH  
 CC AMERICA AND USSR.  
 CC -!- SIMILARITY: BELONGS TO THE BETVI FAMILY OF PATHOGENESIS-RELATED  
 CC PROTEIN.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X77267; CAA54483.1; -;  
 DR HSP; P15494; lBTV.  
 DR InterPro; IPR000916; -;  
 DR Pfam; PF00407; Bet\_v\_1; 1.  
 DR PRINTS; PR00634; BETALLERGEN.  
 DR PROSITE; PS00451; PATHOGENESIS\_BETVI; 1.  
 KW Pathogenesis-related protein; Allergen; Multigene family.  
 FT INIT MET 0  
 SQ SEQUENCE 159 AA; 17316 MW; 3E752543EDD1A08E CRC64;

Query Match 71.4%; Score 5; DB 1; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVGR 6  
 |||||  
 DB 66 YVGR 70

RESULT 7  
 IL7\_HUMAN STANDARD; PRT; 177 AA.  
 ID IL7\_HUMAN  
 AC P13232;  
 DT 01-JAN-1990 (Rel. 13, Created)

```

DR PIR: A25353; A25353.
DR SGD: S0004387; COX8.
KW Oxidoreductase; Mitochondrion; Transit peptide.
FT TRANSIT 1 27 MITOCHONDRION.
FT CHAIN 28 74 CYTOCHROME C OXIDASE POLYPEPTIDE VIII.
FT PROPEP 75 78
FT CONFLICT 47 47 T -> C (IN REF. 3).
FT CONFLICT 68 68 C -> T (IN REF. 3).
SQ SEQUENCE 78 AA; 8907 MW; 61ABE6001BDABDD0 CRC64;

Query Match 71.4%; Score 5; DB 1; Length 78;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7
DB 42 VKGRK 46
|||||

RESULT 2
MP13_FRATU STANDARD; PRT; 123 AA.
AC P18152;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE 13 KDA MAJOR MEMBRANE PROTEIN.
OS Francisella tularensis.
OC Bacteria; Proteobacteria; gamma subdivision; Francisella group;
OC Francisella.
OX NCBI_TaxID=263;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-VACCINE;
RX MEDLINE=90293477; PubMed=1694206;
RA Sjoestedt A., Sandstroem G., Taernvik A., Jaurin B.;
RT "Nucleotide sequence and T cell epitopes of a membrane protein of
Francisella tularensis.";
RL J. Immunol. 145:311-317(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32059; AAA24918.1; -
DR PIR; A37169; A37169.
KW Membrane.
SQ SEQUENCE 123 AA; 13890 MW; 050FCA9BEAD32FC1 CRC64;

Query Match 71.4%; Score 5; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVGRK 6
DB 18 YVGRK 22
|||||

RESULT 3
RS12_RICPR STANDARD; PRT; 129 AA.
AC P41082;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 30S RIBOSOMAL PROTEIN S12.
OS Thermococcus celer.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
OX NCBI_TaxID=2264;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-DSM 2476;
RX MEDLINE=92051399; PubMed=1840672;
RA Klenk H.-P., Schwass V., Zillig W.;
RT "Nucleotide sequence of the genes encoding the L30, S12 and S7
equivalent ribosomal proteins from the archaeum Thermococcus celer.";
RL Nucleic Acids Res. 19:6047-6047(1991).
CC -!- FUNCTION: PROTEIN S12 IS INVOLVED IN THE TRANSLATION INITIATION
CC STEP (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S12P FAMILY OF RIBOSOMAL PROTEINS.

```

```

OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MADRID E;
RC WOOD D.O.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
mitochondria.";
RL Nature 396:133-140(1998).
CC -!- FUNCTION: PROTEIN S12 IS INVOLVED IN THE TRANSLATION INITIATION
CC STEP (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U02603; AAA18329.1; -
DR EMBL; AJ235270; CAA14599.1; -
DR InterPro; IPR000230; -
DR Pfam; PF00164; Ribosomal_S12; 1.
DR PRINTS; PR01034; RIBOSOMALS12.
DR PROSITE; PS00055; RIBOSOMAL_S12; 1.
KW Ribosomal protein.
SQ SEQUENCE 129 AA; 14303 MW; EE7B9A06E28E76F7 CRC64;

Query Match 71.4%; Score 5; DB 1; Length 129;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7
DB 107 VKGRK 111
|||||

RESULT 4
RS12_THECE STANDARD; PRT; 147 AA.
AC P29161;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 30S RIBOSOMAL PROTEIN S12P.
GN RPS12P OR RPS12.
OS Thermococcus celer.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
OX NCBI_TaxID=2264;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-DSM 2476;
RX MEDLINE=92051399; PubMed=1840672;
RA Klenk H.-P., Schwass V., Zillig W.;
RT "Nucleotide sequence of the genes encoding the L30, S12 and S7
equivalent ribosomal proteins from the archaeum Thermococcus celer.";
RL Nucleic Acids Res. 19:6047-6047(1991).
CC -!- FUNCTION: PROTEIN S12 IS INVOLVED IN THE TRANSLATION INITIATION
CC STEP (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S12P FAMILY OF RIBOSOMAL PROTEINS.

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:50:41 : Search time 23.18 seconds  
(without alignments)  
10.345 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_163\_169  
Perfect score: 7  
Sequence: 1 PYVKGRK 7

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	%	Length	DB	ID	Description
1	5	71.4	78	1	COX8_YEAST	P04039	saccharomyc
2	5	71.4	123	1	MP13_FRATU	P18152	francisella
3	5	71.4	129	1	RS12_RICPR	P41082	rickettsia
4	5	71.4	147	1	RS12_THECE	P29161	thermococu
5	5	71.4	150	1	RS12_PYRHO	O59229	pyrococcus
6	5	71.4	159	1	BVE1_BETVE	P43178	betula verr
7	5	71.4	177	1	IL7_HUMAN	P13232	homo sapien
8	5	71.4	238	1	KF5C_RAT	P56536	rattus norv
9	5	71.4	253	1	VATB_METTL	P20022	methanococc
10	5	71.4	325	1	NECD_MOUSE	P25233	mus muscullu
11	5	71.4	342	1	XTNA_CALSA	P23556	caldocellum
12	5	71.4	353	1	OM51_HAEIN	P43840	haemophilus
13	5	71.4	353	1	OM52_HAEIN	P38368	haemophilus
14	5	71.4	359	1	OM53_HAEIN	P45996	haemophilus
15	5	71.4	364	1	YBWL_CAEEL	Q18347	caenorhabdi
16	5	71.4	380	1	COAT_CNV	P15183	cucumber ne
17	5	71.4	391	1	THIL_ZOORA	P07097	zoogloea ra
18	5	71.4	416	1	MURA_BUCAL	P57466	buchnera ap
19	5	71.4	427	1	ATPB_PEPNI	P24467	peptococcus
20	5	71.4	430	1	ACDL_PIG	P79274	scrofa
21	5	71.4	430	1	RT03_MARPO	P26865	marchantia
22	5	71.4	439	1	CYSP_THEPA	Q22497	theileria p
23	5	71.4	460	1	LEU2_LACIA	Q02142	lactococcus
24	5	71.4	464	1	LEU2_SALTY	P15717	salmonella
25	5	71.4	465	1	LEU2_ECOLI	P30127	escherichia
26	5	71.4	467	1	ATPB_PROMO	P29707	prochlorogen
27	5	71.4	470	1	LEU2_AZOVI	P96195	azotobacter
28	5	71.4	477	1	ATPB_STRLI	P50004	streptomyce
29	5	71.4	483	1	ATPB_CORGL	P42464	corynebacte
30	5	71.4	531	1	REF3_BACNO	P39883	bacteroides
31	5	71.4	536	1	ARP_ARATH	P45951	arabidopsis
32	5	71.4	541	1	OPPA_HAEIN	P71370	haemophilus
33	5	71.4	546	1	YTE4_CAEEL	Q17865	caenorhabdi

34	5	71.4	584	1	MAOM_HUMAN	P23368	homo sapien
35	5	71.4	585	1	YNBC_ECOLI	P76092	escherichia
36	5	71.4	588	1	GRN_RAT	P23785	r granulins
37	5	71.4	603	1	PGH2_SHEEP	P79208	ovis aries
38	5	71.4	604	1	PGH2_BOVIN	O62898	bos taurus
39	5	71.4	607	1	UGST_SOLTU	Q00775	solanum tub
40	5	71.4	608	1	UGST_ANTMA	O82627	antirrhinum
41	5	71.4	608	1	UGST_IPOBA	Q42857	ipomoea bat
42	5	71.4	608	1	UGST_MANES	Q43784	manihot esc
43	5	71.4	663	1	Y104_METJA	Q57568	methanococc
44	5	71.4	685	1	AMY1_DICTH	P09961	dictyoglomu
45	5	71.4	688	1	YH10_YEAST	P38708	saccharomyc

#### ALIGNMENTS

RESULT 1  
COX8\_YEAST  
ID COX8\_YEAST STANDARD; PRT; 78 AA.  
AC P04039;  
DT 01-NOV-1986 (Rel. 03, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DE 15-DEC-1998 (Rel. 37, Last annotation update)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE VIII PRECURSOR (EC 1.9.3.1).  
GN COX8 OR YLR395C OR L8084.14.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87057444; PubMed=3023386;  
RA Patterson T.E., Poyton R.O.;  
RT "COX8, the structural gene for yeast cytochrome c oxidase subunit  
RT VIII. DNA sequence and gene disruption indicate that subunit VIII is  
RT required for maximal levels of cellular respiration and is derived  
RT from a precursor which is extended at both its NH2 and COOH  
RT termini.";  
RL J. Biol. Chem. 261:17192-17197(1986).  
[2]  
RP SEQUENCE FROM N.A.  
RN STRAIN=S288C / AB972;  
RC Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,  
RA Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,  
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,  
RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,  
RA Miller N., Nhan M., Pauley A., Peluso D., Rifkin L., Riles L.,  
RA Taich A., Trevasakis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,  
RA Wilson R., Waterson R.;  
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE OF 28-74.  
RN MEDLINE=84212483; PubMed=6327685;  
RC Power S.D., Lochrie M.A., Patterson T.E., Poyton R.O.;  
RT "The nuclear-coded subunits of yeast cytochrome c oxidase. II. The  
RT amino acid sequence of subunit VIII and a model for its disposition  
RT in the inner mitochondrial membrane.";  
RL J. Biol. Chem. 259:6571-6574(1984).  
CC -|- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +  
CC 4 FERRICYTOCHROME C.  
CC -|- SUBUNIT: COMPOSED OF AT LEAST 11 SUBUNITS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J02634; AAA34522.1; -;  
DR EMBL; U19729; AAB82353.1; -;

**THIS PAGE BLANK (USPTO)**



RESULT 12  
F75182  
ribosomal protein S12P PAB0427 - Pyrococcus abyssi (strain Orsay)  
C:Species: Pyrococcus abyssi  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: F75182  
R:anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure  
A:Reference number: A75001  
A:Accession: F75182  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-147 <KAW>  
A:Cross-references: GB:A248284; GB:AL096836; NID:G5457730; PIDN:CAB49541.1; PID:G545805  
A:Experimental source: strain Orsay  
C:Genetics:  
A:Gene: rps12P; PAB0427  
C:Superfamily: Escherichia coli ribosomal protein S12  
F:118/Modified site: beta-methylthioaspartic acid (Asp) #status predicted

Query Match 71.4%; Score 5; DB 2; Length 147;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VKGRK 7  
|||||  
Db 138 VKGRK 142

RESULT 13  
S18713  
ribosomal protein S12 - Thermococcus celer  
C:Species: Thermococcus celer  
C:Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 13-Aug-1999  
C:Accession: S18713; S25568  
R:Klenk, H.P.; Schwass, V.; Zillig, W.  
Nucleic Acids Res. 19, 6047, 1991  
A:Title: Nucleotide sequence of the genes encoding the L30, S12 and S7 equivalent ribosomal proteins  
A:Reference number: S18710; MUID:92051399  
A:Accession: S18713  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-147 <KLE>  
A:Cross-references: EMBL:X60305; NID:G58408; PIDN:CAA42849.1; PID:G58412  
A:Experimental source: DSM 2476  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991  
C:Genetics:  
A:Gene: rps12  
C:Superfamily: Escherichia coli ribosomal protein S12  
C:Keywords: protein biosynthesis; ribosome  
F:118/Modified site: beta-methylthioaspartic acid (Asp) #status predicted

Query Match 71.4%; Score 5; DB 2; Length 147;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VKGRK 7  
|||||  
Db 138 VKGRK 142

RESULT 14  
T02352  
hypothetical protein T8F5.7 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 04-Mar-2000  
C:Accession: T02352  
R:Vysotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Liu, S.; Li, J.; Arau  
Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N.A.; The  
submitted to the EMBL Data Library, July 1998

A:Description: Arabidopsis thaliana chromosome 1 BAC T8F5 complete sequence.  
A:Reference number: Z14666  
A:Accession: T02352  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-148 <VYS>  
A:Cross-references: EMBL:AC004512; NID:G3335331; PID:G3335338; GSPDB:GN00059; ATSP:T8  
C:Genetics:  
A:Gene: ATSP:T8F5.7  
A:Map position: 1  
A:Introns: 28/3; 52/3; 99/1  
C:Superfamily: Arabidopsis thaliana hypothetical protein T8F5.7

Query Match 71.4%; Score 5; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VKGRK 7  
|||||  
Db 127 VKGRK 131

RESULT 15  
D71031  
probable ribosomal protein S12 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
C:Accession: D71031  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A:Reference number: A71000; MUID:98344137  
A:Accession: D71031  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-150 <KAW>  
A:Cross-references: GB:AF000006; NID:G3236133; PIDN:BAA30652.1; PID:G3257969  
A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBa  
C:Genetics:  
A:Gene: PH1542  
C:Superfamily: Escherichia coli ribosomal protein S12  
F:121/Modified site: beta-methylthioaspartic acid (Asp) #status predicted

Query Match 71.4%; Score 5; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VKGRK 7  
|||||  
Db 141 VKGRK 145

Search completed: September 15, 2001, 12:48:55  
Job time: 225 sec

```
RESULT 7
A83185
hypothetical protein PA3688 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: A83185
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bz
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337
A:Accession: A83185
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <STO>
A:Cross-references: GB:AE004788; GB:AE004091; NID:g9949846; PIDN:AAG07076.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3688

Query Match 71.4%; Score 5; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVVKG 5
Db 67 PVVKG 71

RESULT 8
A37169
hypothetical 14K protein (TUL4 5' region) - Francisella tularensis
C:Species: Francisella tularensis
C>Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 04-Mar-2000
C:Accession: A37169
R:Sjoestedt, A.; Sandstroem, G.; Taernvik, A.; Jaurin, B.
J. Immunol. 145, 311-317, 1990
A:Title: Nucleotide sequence and T cell epitopes of a membrane protein of Francisella tu
A:Reference number: A37169; MUID:90293477
A:Accession: A37169
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-123 <SJO>
A:Cross-references: GB:M32059; NID:g148686; PIDN:AAA24918.1; PID:g148687
C:Superfamily: Francisella tularensis hypothetical 14K protein (TUL4 5' region)

Query Match 71.4%; Score 5; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVVKGR 6
Db 18 YVVKGR 22

RESULT 9
H71722
ribosomal protein S12 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: H71722
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499
A:Accession: H71722
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-129 <AND>

A:Cross-references: GB:AJ235270; GB:AJ235269; NID:g3860572; PIDN:CAA14599.1; PID:g386
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: rpsL; RPL30
C:Superfamily: Escherichia coli ribosomal protein S12
F:89/Modified site: beta-methylthioaspartic acid (Asp) #status predicted

Query Match 71.4%; Score 5; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7
Db 107 VKGRK 111

RESULT 10
B32223
interleukin-7 precursor (clone 1) - human
N:Alternate names: IL-7
C:Species: Homo sapiens (man)
C>Date: 20-Oct-1989 #sequence_revision 20-Oct-1989 #text_change 15-Jun-1996
C:Accession: B32223
R:Goodwin, R.G.; Lupton, S.; Schmierer, A.; Hjerrild, K.J.; Jerzy, R.; Clevenger, W.;
Proc. Natl. Acad. Sci. U.S.A. 86, 302-306, 1989
A:Title: Human interleukin 7: molecular cloning and growth factor activity on human a
A:Reference number: A32223; MUID:89098903
A:Accession: B32223
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-133 <GOO>
C:Superfamily: interleukin-7
C:Keywords: cytokine; growth factor

Query Match 71.4%; Score 5; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7
Db 77 VKGRK 81

RESULT 11
C32223
interleukin-7 precursor (clone 4) - human
N:Alternate names: IL-7
C:Species: Homo sapiens (man)
C>Date: 20-Oct-1989 #sequence_revision 20-Oct-1989 #text_change 15-Jun-1996
C:Accession: C32223
R:Goodwin, R.G.; Lupton, S.; Schmierer, A.; Hjerrild, K.J.; Jerzy, R.; Clevenger, W.;
Proc. Natl. Acad. Sci. U.S.A. 86, 302-306, 1989
A:Title: Human interleukin 7: molecular cloning and growth factor activity on human a
A:Reference number: A32223; MUID:89098903
A:Accession: C32223
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-142 <GOO>
C:Superfamily: interleukin-7
C:Keywords: cytokine; growth factor

Query Match 71.4%; Score 5; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VKGRK 7
Db 121 VKGRK 125
```



Query Match 85.7%; Score 6; DB 2; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVGRK 7  
 |||||  
 Db 445 YVGRK 450

RESULT 3  
 C69435  
 hypothetical protein AF1484 - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C:Accession: C69435  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
 A:Reference number: A69250; MUID:98049343  
 A:Accession: C69435  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-53 <KLE>  
 A:Cross-references: GB:AE001000; GB:AE000782; NID:g2689323; PIDN:AA889773.1; PID:g264909

Query Match 71.4%; Score 5; DB 2; Length 53;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVGRK 6  
 |||||  
 Db 18 YVGRK 22

RESULT 4  
 A25353  
 Cytochrome-c oxidase (EC 1.9.3.1) chain VIII precursor - yeast (Saccharomyces cerevisiae  
 N:Alternate names: protein I8084.14; protein YLR395C  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 11-Jun-1999  
 C:Accession: A25353; A05020; S55951  
 R:Patterson, T.E.; Poyton, R.O.  
 J. Biol. Chem. 261, 17192-17197, 1986  
 A:Title: COX8, the structural gene for yeast cytochrome C oxidase subunit VIII.  
 A:Reference number: A25353; MUID:87057444  
 A:Accession: A25353  
 A:Molecule type: DNA  
 A:Residues: 1-78 <PAT>  
 A:Cross-references: GB:J02634; NID:g171296; PIDN:AAA34522.1; PID:g171297  
 R:Power, S.D.; Lochrie, M.A.; Patterson, T.E.; Poyton, R.O.  
 J. Biol. Chem. 259, 6571-6574, 1984  
 A:Title: The nuclear-coded subunits of yeast cytochrome c oxidase: II. the amino acid se  
 A:Reference number: A05020; MUID:84212483  
 A:Accession: A05020  
 A:Molecule type: protein  
 A:Residues: 28-46, 'C', 48-67, 'T', 69-74 <POW>  
 R:Du, Z.  
 submitted to the EMBL Data Library, January 1995  
 A:Description: The sequence of S. cerevisiae cosmid 8084.  
 A:Reference number: S55944  
 A:Accession: S55951  
 A:Molecule type: DNA  
 A:Residues: 1-78 <DUZ>  
 A:Cross-references: EMBL:U19729; NID:9625097; PIDN:AA882353.1; PID:g625111; MIPS:YLR395C  
 C:Genetics:  
 A:Gene: SGD:COX8

A:Cross-references: SGD:S0004387; MIPS:YLR395C  
 A:Map position: 12R  
 A:Genome: nuclear  
 C:Superfamily: cytochrome-c oxidase chain VIIc  
 C:Keywords: membrane-associated complex; mitochondrion; oxidative phosphorylation; ox  
 F:1-71/Domain: transit peptide (mitochondrion) #status experimental <TNP>  
 F:28-78/Product: cytochrome-c oxidase chain VIII #status experimental <MAT>

Query Match 71.4%; Score 5; DB 2; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YVGRK 7  
 |||||  
 Db 42 YVGRK 46

RESULT 5  
 S20774  
 Ig heavy chain V region - human  
 C:Species: Homo sapiens (man)  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Sep-1997  
 C:Accession: S20774  
 R:Mortari, F.; Wang, J.; Schroeder, H.W.  
 submitted to the EMBL Data Library, April 1992  
 A:Description: Analysis of human cord blood Ig heavy chain IgA and IgG repertoire.  
 A:Reference number: S20764  
 A:Accession: S20774  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-106 <MOR>  
 A:Cross-references: EMBL:Z11942; NID:g33880; PID:g33881  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 71.4%; Score 5; DB 2; Length 106;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVGRK 6  
 |||||  
 Db 45 YVGRK 49

RESULT 6  
 PH1659  
 Ig heavy chain V region (clone KL1) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 16-Aug-1996  
 C:Accession: PH1659  
 R:Hillson, J.L.; Karr, N.S.; Opliger, I.R.; Mannik, M.; Sasso, E.H.  
 J. Exp. Med. 178, 331-336, 1993  
 A:Title: The structural basis of germline-encoded VH3 immunoglobulin binding to staph  
 A:Reference number: PH1642; MUID:93301610  
 A:Accession: PH1659  
 A:Molecule type: mRNA  
 A:Residues: 1-111 <HIL>  
 A:Experimental source: B cell  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:7-88/Domain: immunoglobulin homology <IMM>

Query Match 71.4%; Score 5; DB 2; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVGRK 6  
 |||||  
 Db 53 YVGRK 57

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:48:54 ; Search time 45.39 Seconds  
(without alignments)  
11.748 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_163\_169

Perfect score: 7

Sequence: 1 PYVKGRK 7

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR\_68.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	85.7	471	2 T26612	hypothetical prote
2	6	85.7	487	2 T32635	hypothetical prote
3	5	71.4	53	2 C69435	hypothetical prote
4	5	71.4	78	2 A25353	cytochrome-c oxida
5	5	71.4	106	2 S20774	Ig heavy chain V r
6	5	71.4	111	2 PH1659	Ig heavy chain V r
7	5	71.4	114	2 A83185	hypothetical prote
8	5	71.4	123	2 A37169	ribosomal protein
9	5	71.4	129	2 H71722	hypothetical 14K p
10	5	71.4	133	2 B32223	interleukin-7 prec
11	5	71.4	142	2 C32223	interleukin-7 prec
12	5	71.4	147	2 F75182	ribosomal protein
13	5	71.4	147	2 S18713	ribosomal protein
14	5	71.4	148	2 T02352	hypothetical prote
15	5	71.4	150	2 D71031	probable ribosomal
16	5	71.4	154	2 P00465	beta C protein - p
17	5	71.4	160	2 D55699	major pollen aller
18	5	71.4	163	2 A86639	transcription regu
19	5	71.4	177	2 A32223	interleukin-7 prec
20	5	71.4	187	2 T07802	ADPglucose--starch
21	5	71.4	187	2 T07804	ADPglucose--starch
22	5	71.4	194	2 C70209	conserved hypothe
23	5	71.4	197	2 D70017	hypothetical prote
24	5	71.4	235	2 S60156	chlamyopsin - Chla
25	5	71.4	252	2 T29916	hypothetical prote
26	5	71.4	253	2 S05029	H+-transporting AT
27	5	71.4	256	2 S42584	hypothetical prote
28	5	71.4	278	2 S35003	probable transposa
29	5	71.4	278	2 S35004	probable transposa

## ALIGNMENTS

### RESULT 1

T26612

hypothetical protein Y37AlB.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000

C:Accession: T26612

R:McMurray, A.

submitted to the EMBL Data Library, June 1998

A:Reference number: Z20245

A:Accession: T26612

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-471 <WIL>

A:Cross-references: EMBL:AL023835; PIDN:CAA13490.1; GSPDB:GN00022; CESP:Y37AlB.5

A:Experimental source: clone Y37AlB

C:Genetics:

A:Gene: CESP:Y37AlB.5

A:Map position: 4

A:Introns: 58/3; 91/2; 215/1; 282/3; 418/2

C:Superfamily: Caenorhabditis elegans hypothetical protein Y37AlB.5

Query Match 85.7%; Score 6; DB 2; Length 471;  
Best Local Similarity 100.0%; Pred.No. 12;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVKGRK 7

|||||

Db 376 YVKGRK 381

### RESULT 2

T32635

hypothetical protein F42G8.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 02-Jun-2000

C:Accession: T32635

R:Gattung, S.; Holmes, A.

submitted to the EMBL Data Library, December 1997

A:Description: The sequence of C. elegans cosmid F42G8.

A:Reference number: Z21203

A:Accession: T32635

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-487 <GAV>

A:Cross-references: EMBL:AF038618; PIDN:AAB92072.1; GSPDB:GN00022; CESP:F42G8.8

A:Experimental source: strain Bristol N2; clone F42G8

C:Genetics:

A:Gene: CESP:F42G8.8

A:Map position: 4

A:Introns: 27/3; 97/2; 122/3; 203/3; 236/2; 274/3; 302/3; 377/1

C:Superfamily: phosphoprotein phosphatase; phosphoesterase core homology; phosphoprot

probable transposa  
hypothetical prote  
hypothetical prote  
ABC transporter -  
RNA-directed DNA p  
necdin, brain - mo  
probable transposa  
endo-1.4-beta-xyla  
probable transcrip  
hypothetical prote  
outer membrane pro  
hypothetical prote  
hypothetical prote  
transcription fact  
chitinase (EC 3.2.  
coat protein - cuc

**THIS PAGE BLANK (USPTO)**

RESULT 15  
5229279-2  
; Patent No. 5229279  
; APPLICANT: PEOPLES, OLIVER P.; SINSKEY, ANTHONY J.  
; TITLE OF INVENTION: METHOD FOR PRODUCING NOVEL POLYESTER  
; BIOPOLYMERS  
; NUMBER OF SEQUENCES: 8  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/556,535  
; FILING DATE: 13-AUG-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 67,695  
; FILING DATE: 29-AUG-1987  
; SEQ ID NO: 2:  
; LENGTH: 391  
5229279-2

Query Match 71.4%; Score 5; DB 6; Length 391;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 VKGRK 7  
Db 203 VKGRK 207

Search completed: September 15, 2001, 12:47:56  
Job time: 186 sec

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (216) 622-8458  
TELEFAX: (216) 241-0816  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 359 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-467-722A-2

Query Match 71.4%; Score 5; DB 3; Length 359;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YKGRK 7  
DB 335 YKGRK 339  
|||||

RESULT 13  
US-08-591-629-2  
; Sequence 2, Application US/08591629  
; Patent No. 5993808  
; GENERAL INFORMATION:  
; APPLICANT: MELCHERS, Leo Sjoerd  
; APPLICANT: APOTHEKER-DE GROOT, Marion  
; APPLICANT: BOL, John Ferdinand  
; APPLICANT: CORNELISSEN, Bernardus Johannes Clemens  
; APPLICANT: LINTHORST, Hubertus Josephus Maria  
; APPLICANT: PONSTEIN, Anne Silene  
; APPLICANT: SELA-BUURLAGE, Marianne Beatrix  
; TITLE OF INVENTION: Plant chitinases, DNA coding therefor and  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ladas & Parry  
; STREET: 26 West 61st Street  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10023-7604  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
; COMPUTER: IBM PC 4.86 SX 50 Mhz  
; OPERATING SYSTEM: DOS 6.20  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/591.629  
; FILING DATE: 15-FEB-96  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP94/02761  
; FILING DATE: 17-AUG-94  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 93202425.0  
; FILING DATE: 17-AUG-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MASS, CLIFFORD J.  
; REGISTRATION NUMBER: 30,086  
; REFERENCE/DOCKET NUMBER: U-010627-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 708-1800  
; TELEFAX: (212) 246-8959  
; TELEX: 233288  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 371 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-591-629-2

Query Match 71.4%; Score 5; DB 2; Length 371;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVKGK 6  
DB 333 YVKGK 337  
|||||

RESULT 14  
US-08-591-629-8  
; Sequence 8, Application US/08591629  
; Patent No. 5993808  
; GENERAL INFORMATION:  
; APPLICANT: MELCHERS, Leo Sjoerd  
; APPLICANT: APOTHEKER-DE GROOT, Marion  
; APPLICANT: BOL, John Ferdinand  
; APPLICANT: CORNELISSEN, Bernardus Johannes Clemens  
; APPLICANT: LINTHORST, Hubertus Josephus Maria  
; APPLICANT: PONSTEIN, Anne Silene  
; APPLICANT: SELA-BUURLAGE, Marianne Beatrix  
; TITLE OF INVENTION: Plant chitinases, DNA coding therefor and  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ladas & Parry  
; STREET: 26 West 61st Street  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10023-7604  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
; COMPUTER: IBM PC 4.86 SX 50 Mhz  
; OPERATING SYSTEM: DOS 6.20  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/591.629  
; FILING DATE: 15-FEB-96  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP94/02761  
; FILING DATE: 17-AUG-94  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 93202425.0  
; FILING DATE: 17-AUG-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MASS, CLIFFORD J.  
; REGISTRATION NUMBER: 30,086  
; REFERENCE/DOCKET NUMBER: U-010627-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 708-1800  
; TELEFAX: (212) 246-8959  
; TELEX: 233288  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 377 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-591-629-8

Query Match 71.4%; Score 5; DB 2; Length 377;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YVKGK 6  
DB 339 YVKGK 343  
|||||

STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1040691  
US-08-773-870-6

Query Match 71.4%; Score 5; DB 2; Length 325;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 YVGR 6  
|||||  
Db 203 YVGR 207

RESULT 10  
US-08-282-197C-57  
Sequence 57, Application US/08282197C  
Patent No. 5871730  
GENERAL INFORMATION:  
APPLICANT: Brzezinski, Ryszard  
APPLICANT: Dery, Claude V  
APPLICANT: Beaulieu, Carole  
TITLE OF INVENTION: Thermostable Xylanase DNA, Protein and  
METHODS OF USE  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Ave., NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/282.197C  
FILING DATE: 29-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Cimbala, Michele A  
REGISTRATION NUMBER: 33,851  
REFERENCE/DOCKET NUMBER: 1050.0410000  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 343 amino acids  
TYPE: amino acid  
STRANDEDNESS: both  
TOPOLOGY: both  
US-08-282-197C-57

Query Match 71.4%; Score 5; DB 2; Length 343;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VVGR 7  
|||||  
Db 316 VVGR 320

RESULT 11  
US-08-457-997B-2  
Sequence 2, Application US/08457997B  
Patent No. 5766608

GENERAL INFORMATION:  
APPLICANT: Kolattukudy, P. E.  
TITLE OF INVENTION: Otitis Media Vaccine  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Calfee, Halter and Griswold  
STREET: Suite 1800 800 Superior Avenue  
CITY: Cleveland  
STATE: Ohio  
COUNTRY: U.S.A.  
ZIP: 44114-2688  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,997B  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Gollrick, Mary E.  
REGISTRATION NUMBER: 34,829  
REFERENCE/DOCKET NUMBER: 22727/00102  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (216) 622-8458  
TELEFAX: (216) 241-0816  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 359 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-457-997B-2

Query Match 71.4%; Score 5; DB 1; Length 359;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VVGR 7  
|||||  
Db 335 VVGR 339

RESULT 12  
US-08-467-722A-2  
Sequence 2, Application US/08467722A  
Patent No. 6030626  
GENERAL INFORMATION:  
APPLICANT: Kolattukudy, P. E.  
TITLE OF INVENTION: Otitis Media Vaccine  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Calfee, Halter and Griswold  
STREET: Suite 1800 800 Superior Avenue  
CITY: Cleveland  
STATE: Ohio  
COUNTRY: U.S.A.  
ZIP: 44114-2688  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,722A  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Gollrick, Mary E.  
REGISTRATION NUMBER: 34,829  
REFERENCE/DOCKET NUMBER: 22727/00102

```

; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Apple 7.1
; SOFTWARE: Microsoft Word, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,161
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,908
; FILING DATE: 22-MAY-1995
; APPLICATION NUMBER: US 08/231,205
; FILING DATE: 21-APR-1994
; APPLICATION NUMBER: US 07/957,649
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,438
; FILING DATE: 13-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/255,209
; FILING DATE: 07-OCT-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/113,566
; FILING DATE: 26-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2104-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-871-161-4

```

```

Query Match 71.4%; Score 5; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 VKGRK 7
DB 121 VKGRK 125

```

```

RESULT 8
PCT-US95-08950-7
; SEQUENCE 7, Application PC/TUS9508950
; GENERAL INFORMATION:
; APPLICANT: Zurawski, Sandra M.
; APPLICANT: Zurawski, Gerard
; TITLE OF INVENTION: MUTEINS OF MAMMALIAN CYTOKINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08950
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/284,393
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0389
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-852-9196
; TELEFAX: 415-496-1200
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-08950-7

```

```

Query Match 71.4%; Score 5; DB 5; Length 177;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 VKGRK 7
DB 121 VKGRK 125

```

```

RESULT 9
US-08-773-870-6
; SEQUENCE 6, Application US/08773870
; Patent No. 5912143
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN MAGE-LIKE PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/773,870
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0179 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid

```

Sequence 4, Application US/08446908  
Patent No. 5705149  
GENERAL INFORMATION:  
APPLICANT: Namen, Anthony E.  
APPLICANT: Goodwin, Raymond G.  
APPLICANT: Lupton, Stephen D.  
APPLICANT: Mochizuki, Diane Y.  
TITLE OF INVENTION: Interleukin-7 and Antibodies Reactive  
TITLE OF INVENTION: Therewith  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: US  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Apple 7.1  
SOFTWARE: Microsoft Word, Version 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,908  
FILING DATE: 22-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/231,205  
FILING DATE: 21-APR-1994  
APPLICATION NUMBER: US 07/957,649  
FILING DATE: 06-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/511,438  
FILING DATE: 13-APR-1990  
APPLICATION NUMBER: US 07/255,209  
FILING DATE: 07-OCT-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/113,566  
FILING DATE: 26-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Seese, Kathryn A.  
REGISTRATION NUMBER: 32,172  
REFERENCE/DOCKET NUMBER: 2104-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 177 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-446-908-4

Query Match 71.4%; Score 5; DB 1; Length 177;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 VKGRK 7  
Db 121 VKGRK 125

RESULT 6  
US-08-231-205A-4  
Sequence 4, Application US/08231205A  
Patent No. 5714585  
GENERAL INFORMATION:  
APPLICANT: Namen, Anthony E.  
APPLICANT: Goodwin, Raymond G.  
APPLICANT: Lupton, Stephen D.

APPLICANT: Mochizuki, Diane Y.  
TITLE OF INVENTION: Interleukin-7 and Antibodies Reactive  
TITLE OF INVENTION: Therewith  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: US  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Apple 7.1  
SOFTWARE: Microsoft Word, Version 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,205A  
FILING DATE: 21-APR-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/957,649  
FILING DATE: 06-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/511,438  
FILING DATE: 13-APR-1990  
APPLICATION NUMBER: US 07/255,209  
FILING DATE: 07-OCT-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/113,566  
FILING DATE: 26-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Seese, Kathryn A.  
REGISTRATION NUMBER: 32,172  
REFERENCE/DOCKET NUMBER: 2104-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 177 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-231-205A-4

Query Match 71.4%; Score 5; DB 1; Length 177;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 VKGRK 7  
Db 121 VKGRK 125

RESULT 7  
US-08-871-161-4  
Sequence 4, Application US/08871161  
Patent No. 5965122  
GENERAL INFORMATION:  
APPLICANT: Namen, Anthony E.  
APPLICANT: Goodwin, Raymond G.  
APPLICANT: Lupton, Stephen D.  
APPLICANT: Mochizuki, Diane Y.  
TITLE OF INVENTION: Interleukin-7 and Antibodies Reactive  
TITLE OF INVENTION: Therewith  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle



US-08-470-720-13

Query Match 71.4%; Score 5; DB 2; Length 79;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VKGRK 7  
Db 3 VKGRK 7

RESULT 2

Patent No. 5229115  
; APPLICANT: LYNCH, DAVID H.  
; TITLE OF INVENTION: ADOPTIVE IMMUNOTHERAPY WITH INTERLEUKIN-7  
; NUMBER OF SEQUENCES: 2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/559,001  
; FILING DATE: 26-JUL-1990  
; SEQ ID NO:1:  
; LENGTH: 151  
5229115-1

Query Match 71.4%; Score 5; DB 6; Length 151;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VKGRK 7  
Db 96 VKGRK 100

RESULT 3

US-08-318-193-84  
; Sequence 84, Application US/08318193  
; Patent No. 5641663  
; GENERAL INFORMATION:  
; APPLICANT: GARVIN, Robert T.  
; APPLICANT: MALEK, Lawrence T.  
; TITLE OF INVENTION: AN EXPRESSION SYSTEM FOR THE SECRETION  
; TITLE OF INVENTION: OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY  
; TITLE OF INVENTION: STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS  
; TITLE OF INVENTION: PROTEINS FROM STREPTOMYCES  
; NUMBER OF SEQUENCES: 91  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/318,193  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,314  
; FILING DATE:  
; APPLICATION NUMBER: US 07/224,568  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 18740/116 CACO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109  
TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 84:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 152 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-318-193-84

Query Match 71.4%; Score 5; DB 1; Length 152;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VKGRK 7  
Db 96 VKGRK 100

RESULT 4

US-08-284-393B-7  
; Sequence 7, Application US/08284393B  
; Patent No. 5696234  
; GENERAL INFORMATION:  
; APPLICANT: Zurawski, Sandra M.  
; APPLICANT: Zurawski, Gerard  
; TITLE OF INVENTION: MUTEINS OF MAMMALIAN CYTOKINES  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DNAX Research Institute  
; STREET: 901 California Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/284,393B  
; FILING DATE: 01-AUG-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ching, Edwin P.  
; REGISTRATION NUMBER: 34,090  
; REFERENCE/DOCKET NUMBER: DX0389  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-852-9196  
; TELEFAX: 415-496-1200  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 177 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-284-393B-7

Query Match 71.4%; Score 5; DB 1; Length 177;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VKGRK 7  
Db 121 VKGRK 125

RESULT 5

US-08-446-908-4

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: September 15, 2001, 12:47:54 ; Search time 35,36 Seconds  
(without alignments)  
4.076 Million cell updates/sec

Title: us-09-389-000-2\_COPY\_163\_169

Perfect score: 7  
Sequence: 1 PPKGRK 7

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents\_AA.\*  
1: /cgnl\_7/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgnl\_7/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgnl\_7/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgnl\_7/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgnl\_7/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgnl\_7/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	71.4	79	2	US-08-470-720-13
2	5	71.4	151	6	5229115-1
3	5	71.4	152	1	US-08-318-193-84
4	5	71.4	177	1	US-08-284-393B-7
5	5	71.4	177	1	US-08-446-908-4
6	5	71.4	177	1	US-08-231-205A-4
7	5	71.4	177	2	US-08-871-161-4
8	5	71.4	177	5	PCT-US95-08950-7
9	5	71.4	325	2	US-08-773-870-6
10	5	71.4	343	2	US-08-282-197C-57
11	5	71.4	359	1	US-08-457-997B-2
12	5	71.4	359	3	US-08-467-722A-2
13	5	71.4	371	2	US-08-591-629-2
14	5	71.4	377	2	US-08-591-629-8
15	5	71.4	391	6	5229279-2
16	5	71.4	391	6	5512669-2
17	5	71.4	411	2	US-08-713-815A-4
18	5	71.4	441	2	US-08-713-815A-3
19	5	71.4	460	1	US-08-403-866-3
20	5	71.4	531	4	US-08-984-618-8
21	5	71.4	589	1	US-07-668-648-2
22	5	71.4	589	2	US-08-429-998-2
23	5	71.4	589	2	US-08-431-333-2
24	5	71.4	589	5	PCT-US91-02321-2
25	5	71.4	1019	1	US-08-271-364A-7
26	5	71.4	1019	2	US-08-222-715B-26
27	5	71.4	1068	3	US-08-390-874C-11

28	5	71.4	1069	2	US-08-162-081B-37	Sequence 37, Appl
29	5	71.4	1069	2	US-08-780-872-37	Sequence 37, Appl
30	5	71.4	1080	2	US-08-162-081B-36	Sequence 36, Appl
31	5	71.4	1080	2	US-08-780-872-36	Sequence 36, Appl
32	4	57.1	5	1	US-08-843-035-9	Sequence 9, Appl
33	4	57.1	5	2	US-08-755-496A-9	Sequence 9, Appl
34	4	57.1	5	3	US-08-813-586-20	Sequence 20, Appl
35	4	57.1	6	1	US-08-843-035-10	Sequence 10, Appl
36	4	57.1	6	2	US-08-755-496A-10	Sequence 10, Appl
37	4	57.1	7	1	US-08-843-035-11	Sequence 11, Appl
38	4	57.1	7	2	US-08-755-496A-11	Sequence 11, Appl
39	4	57.1	8	1	US-08-843-035-12	Sequence 12, Appl
40	4	57.1	8	2	US-08-755-496A-12	Sequence 12, Appl
41	4	57.1	9	1	US-08-030-077-4	Sequence 4, Appl
42	4	57.1	9	1	US-08-843-035-13	Sequence 13, Appl
43	4	57.1	9	1	US-08-843-035-36	Sequence 13, Appl
44	4	57.1	9	2	US-08-755-496A-13	Sequence 13, Appl
45	4	57.1	9	2	US-08-755-496A-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1  
US-08-470-720-13  
; Sequence 13, Application US/08470720  
; Patent No. 5824798  
; GENERAL INFORMATION:  
; APPLICANT: HOFVANDER, Per  
; APPLICANT: PERSSON, Per T  
; APPLICANT: WIKSTROM, Olle  
; APPLICANT: TALLBERG, Anneli  
; TITLE OF INVENTION: GENETICALLY ENGINEERED MODIFICATION OF  
; TITLE OF INVENTION: POTATO TO FORM AMYLOPECTIN-TYPE STARCH  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: George Mason Bldg., Washington & Prince Sts.  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/470,720  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/070,455  
; FILING DATE: 09-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crane-Feury, Sharon E  
; REGISTRATION NUMBER: 36,113  
; REFERENCE/DOCKET NUMBER: 003300-293  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 79 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 1..79  
; OTHER INFORMATION: /note= "Amino acid sequence encoded  
; by nucleotides 1349-1585 of SEQ ID NO 2."

Oy 7 LRRERI 12  
| | | | |  
Db 15 LRRERI 20

Search completed: September 15, 2001, 12:43:04  
Job time: 430 sec

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*,"  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003615; AAF52458.2;  
 DR FlyBase; FBgn0031869; CG18304.  
 DR InterPro; IPR002017;  
 KW Hypothetical protein.  
 SQ SEQUENCE 1833 AA; 206011 MW; ED1948482573584E CRC64;

Query Match 46.7%; Score 7; DB 5; Length 1833;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EKLRRER 11

|||||

Db 1479 EKLRRER 1485

RESULT 13

QSUBN9

ID Q9UBN9 PRELIMINARY; PRT; 39 AA.

AC Q9UBN9;

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE E6-AP UBIQUITIN-PROTEIN LIGASE (FRAGMENT).

GN UBE3A.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RX MEDLINE=98126441; PubMed=9465301;

RA Kishino T., Wagstaff J.;

RT "Genomic organization of the UBE3A/E6-AP gene and related

RT pseudogenes,"

RL Genomics 47:101-107(1998).

RN [2]

RP SEQUENCE FROM N.A.

RA Hennies H.C., Buerger J., Sperling K., Reis A.;

RT "Mutations in the E6-AP gene (UBE3A) in patients with Angelman

RT syndrome,"

RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF009341; AAC39580.1;

DR EMBL; AJ001113; CAA04540.1;

DR InterPro; IPR000569;

DR PROSITE; PS02037; HECT. 1.

KW Ligase.

FT NON\_TER

SQ SEQUENCE 39 AA; 4441 MW; 763722F374FA7193 CRC64;

Query Match

Best Local Similarity 40.0%; Score 6; DB 4; Length 39;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKKEL 7

|||||

Db 17 SSKKEL 22

RESULT 14

Q9EY25

ID Q9EY25 PRELIMINARY; PRT; 64 AA.

AC Q9EY25;

DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE HOST FACTOR I (FRAGMENT).

OS Photobacterium profundum (Photobacterium sp. (strain SS9)).

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae;

OC Photobacterium.

OX NCBI\_TaxID=74109;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SS9;

RA Bidle K.A., Bartlett D.H.;

RT "An RNA arbitrarily primed PCR survey of genes regulated at low and

RT high pressure by ToxR in the marine bacterium *Photobacterium profundum*

RT SS9,"

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF307979; AAG34558.1;

FT NON\_TER 64

SQ SEQUENCE 64 AA; 7345 MW; 5BF64E6B23F24FB9 CRC64;

Query Match

Best Local Similarity 40.0%; Score 6; DB 2; Length 64;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERI 12

|||||

Db 15 LRRERI 20

RESULT 15

Q9KV11

ID Q9KV11 PRELIMINARY; PRT; 87 AA.

AC Q9KV11;

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE HOST FACTOR-I, PUTATIVE.

GN VC0347.

OS Vibrio cholerae.

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

OX NCBI\_TaxID=666;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=EL TOR N16961 / SEROTYPE O1;

RX MEDLINE=20406833; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,

RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,

RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,

RA Encolleva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,

RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,

RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,

RA Fraser C.M.;

RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*

RT cholerae,"

RL Nature 406:477-483(2000).

DR EMBL; AE004123; AAF93520.1;

DR TIGR; VC0347;

DR InterPro; IPR000774;

DR PRODOM; PD001516;

SQ SEQUENCE 87 AA; 9769 MW; A7E4B6878A164DB0 CRC64;

Query Match

Best Local Similarity 40.0%; Score 6; DB 2; Length 87;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SQ SEQUENCE 1260 AA; 143958 MW; 2074957B2146DE06 CRC64;

Query Match 46.7%; Score 7; DB 5; Length 1260;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KEKLRE 10  
 DB 887 KEKLRE 893

RESULT 10  
 Q08281 ID Q08281 PRELIMINARY; PRT; 1341 AA.  
 AC Q08281;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE CHROMOSOME XV READING FRAME ORF YOL138C.  
 GN YOL138C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Arino J., Casamayor A., Gamo F.J., Gancedo C., Lafuente M.J.,  
 RA Aldea M., Casas C., Herrero E.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: CONTAINS A RING (C3HC4-CLASS) ZINC FINGER.  
 DR EMBL; Z74880; CAA99159.1; -;  
 DR SGD; S0005498; YOL138C.  
 DR InterPro; IPR001860; -;  
 DR InterPro; IPR001841; -;  
 DR InterPro; IPR002052; -;  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR Pfam; PF00400; WD40; 2.  
 DR PROSITE; PS00092; N6\_MTASE; UNKNOWN\_1.  
 DR PROSITE; PS00678; WD\_REPEATS; UNKNOWN\_1.  
 DR SMART; SM00184; RING; 1.  
 KW Repeat; WD repeat; Zinc-finger.  
 SQ SEQUENCE 1341 AA; 149268 MW; EB0EEAD034818EA0 CRC64;

Query Match 46.7%; Score 7; DB 3; Length 1341;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKELR 8  
 DB 1257 SSKELR 1263

RESULT 11  
 Q92271 ID Q92271 PRELIMINARY; PRT; 1342 AA.  
 AC Q92271;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE 12.8 KBP FRAGMENT OF THE LEFT ARM OF CHROMOSOME XV.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FY1679;

RX MEDLINE=97051593; PubMed=8896270;  
 RA Aldea M., Piedrafita L., Casas C., Casamayor A., Khalid H.,  
 RA Balcells L., Arino J., Herrero E.;  
 RT "Sequence analysis of a 12 801 bp fragment of the left arm of yeast  
 chromosome XV containing a putative 6-phosphofructo-2-kinase gene, a  
 gene for a possible glycopospholipid-anchored surface protein and six  
 other open reading frames";  
 RL Yeast 12:1053-1058(1996).  
 CC -!- SIMILARITY: CONTAINS A RING (C3HC4-CLASS) ZINC FINGER.  
 DR EMBL; X95465; CAA64732.1; -;  
 DR InterPro; IPR001680; -;  
 DR InterPro; IPR001841; -;  
 DR InterPro; IPR002052; -;  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR Pfam; PF00400; WD40; 2.  
 DR PROSITE; PS00092; N6\_MTASE; UNKNOWN\_1.  
 DR PROSITE; PS00678; WD\_REPEATS; UNKNOWN\_1.  
 DR SMART; SM00184; RING; 1.  
 KW Repeat; WD repeat; Zinc-finger.  
 SQ SEQUENCE 1342 AA; 149367 MW; E874142A08D83E1B CRC64;

Query Match 46.7%; Score 7; DB 3; Length 1342;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKELR 8  
 DB 1258 SSKELR 1264

RESULT 12  
 Q9VM67 ID Q9VM67 PRELIMINARY; PRT; 1833 AA.  
 AC Q9VM67; Q9VM66;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE CGI8304 PROTEIN.  
 GN CGI8304 OR CGI3774.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry J., Morris J., Moshrefi A.,

ID O45358 PRELIMINARY; PRT; 791 AA.  
 AC O45358; O45404;  
 DT 01-JUN-1998 (TremBLrel. 06, Created)  
 DT 01-NOV-1998 (TremBLrel. 08, Last sequence update)  
 DT 01-JUN-2000 (TremBLrel. 14, Last annotation update)  
 DE F26D2.2 PROTEIN.  
 GN F26D2.2  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RL McMurray A.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
 RA Smaldon N., Smith A., Sonnenhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA McMurray A.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z93377; CAB07581.1; -;  
 DR EMBL; Z81513; CAB07581.1; JOINED.  
 DR EMBL; Z81513; CAB04186.1; -;  
 DR EMBL; Z93377; CAB04186.1; JOINED.  
 SQ SEQUENCE 791 AA; 91505 MW; 9B53803A9B6363F0 CRC64;  
  
 Query Match 46.7%; Score 7; DB 5; Length 791;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 5 EKLRRER 11  
 Db 153 EKLRRER 159  
 |||||  
 |||||  
  
 RESULT 7  
 Q9WVP5 PRELIMINARY; PRT; 1080 AA.  
 ID Q9WVP5  
 AC Q9WVP5;  
 DT 01-NOV-1999 (TremBLrel. 12, Created)  
 DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)  
 DE PUTATIVE DEUBIQUITINATING ENZYME UBPY.  
 GN UBPY.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Martegani E., Gnesutta N., Mauri I., Bini V.;  
 RT "Mouse deubiquitinating enzyme (m-ubpy).";  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF057146; AAD38869.1; -;  
 DR MEROPS; C19.011; -;  
 DR InterPro; IPR001394; -;  
 DR InterPro; IPR001448; -;

DR InterPro; IPR001763; -;  
 DR Pfam; PF00442; UCH-1; 1.  
 DR Pfam; PF00443; UCH-2; 1.  
 DR Pfam; PF00581; Rhodanese; 1.  
 DR PROSITE; PS00304; SASP\_1; UNKNOWN\_1.  
 DR PROSITE; PS00972; UCH\_2\_1; 1.  
 DR PROSITE; PS00973; UCH\_2\_2; 1.  
 DR PROSITE; PS0235; UCH\_2\_3; 1.  
 DR SMART; SM00450; RHOD; 1.  
 SQ SEQUENCE 1080 AA; 122578 MW; 2BF00AFF68AFC8B5 CRC64;  
  
 Query Match 46.7%; Score 7; DB 11; Length 1080;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 4 KEKLRR 10  
 Db 485 KEKLRR 491  
 |||||  
 |||||  
  
 RESULT 8  
 Q9EQU1 PRELIMINARY; PRT; 1080 AA.  
 ID Q9EQU1  
 AC Q9EQU1;  
 DT 01-MAR-2001 (TremBLrel. 16, Created)  
 DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)  
 DE DEUBIQUITINATING ENZYME UBPY.  
 GN MURPY.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20545486; PubMed=10982817;  
 RA Kato M., Miyazawa K., Kitamura N.;  
 RT "A Deubiquitinating Enzyme UBPY Interacts with the Src Homology 3  
 RT Domain of Hrs-binding Protein via a Novel Binding Motif  
 RT PX(V/I)(D/N)RXKP.";  
 RL J. Biol. Chem. 275:37481-37487(2000).  
 DR EMBL; AB045709; BAB18534.1; -;  
 SQ SEQUENCE 1080 AA; 122460 MW; C1BB770FF98C3224 CRC64;  
  
 Query Match 46.7%; Score 7; DB 11; Length 1080;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 4 KEKLRR 10  
 Db 485 KEKLRR 491  
 |||||  
 |||||  
  
 RESULT 9  
 Q9GRG5 PRELIMINARY; PRT; 1260 AA.  
 ID Q9GRG5  
 AC Q9GRG5;  
 DT 01-MAR-2001 (TremBLrel. 16, Created)  
 DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)  
 DE PUTATIVE STRUCTURAL MAINTENANCE OF CHROMOSOME 3 PROTEIN.  
 GN PUTATIVE SMC3.  
 OS Trypanosoma brucei.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 OX NCBI\_TaxID=5691;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ersfeld K., Walgraffe D., Gull K.;  
 RT "Identification of an SMC3-homologue in Trypanosoma brucei.";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ279087; CAC12695.1; -;

GN AT4G05400.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,  
 RA Preston R.A., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,  
 RA Shekher M., Schütz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,  
 RA Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL161503; CAB81082.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 250 AA; 27627 MW; 7AAA04FBF0A6F707 CRC64;

Query Match 46.7%; Score 7; DB 10; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSKELR 8  
 |||||  
 Db 244 SSKELR 250

RESULT 3  
 ID Q9LYR0 PRELIMINARY; PRT; 443 AA.  
 AC Q9LYR0;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE HYPOTHETICAL 50.8 KDA PROTEIN.  
 GN T22N19.110.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,  
 RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL163572; CAB87153.1; -;  
 DR InterPro; IPR000048; -;  
 DR Pfam; PF00612; IQ; 2.  
 DR SMART; SM00015; IQ; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 443 AA; 50820 MW; EA3CB8B5E2041D83 CRC64;

Query Match 46.7%; Score 7; DB 10; Length 443;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRERIK 13  
 |||||  
 Db 227 LRRERIK 233

RESULT 4  
 Q9ZB31

ID Q9ZB31 PRELIMINARY; PRT; 502 AA.  
 AC Q9ZB31;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE ERYTHRITOL PHOSPHATE DEHYDROGENASE.  
 GN ERIB.  
 OS Brucella abortus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=235;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=2308;  
 RA Sangari F.J., Aguero J., Garcia-Lobo J.M.;  
 RT "Characterization of a gene cluster involved in erythritol catabolism  
 in Brucella abortus."  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U57100; AAD11520.1; -;  
 DR InterPro; IPR000447; -;  
 DR Pfam; PF01224; FAD\_Gly3P\_dh; 1.  
 DR PRINTS; PR01001; FADG3PDH.  
 DR PROSITE; PS00977; FAD\_G3PDH\_1; 1.  
 SQ SEQUENCE 502 AA; 56231 MW; A88459636F4D7A11 CRC64;

Query Match 46.7%; Score 7; DB 2; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KEKLRE 10  
 |||||  
 Db 307 KEKLRE 313

RESULT 5  
 ID Q9FNF6 PRELIMINARY; PRT; 746 AA.  
 AC Q9FNF6;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE DIPEPTIDYL PEPTIDASE IV-LIKE PROTEIN.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=98069011; PubMed=9405937;  
 RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.  
 Sequence features of the regions of 1,044,062 bp covered by thirteen  
 physically assigned P1 clones."  
 RL DNA Res. 4:291-300(1997).  
 DR EMBL: AB006701; BAB10391.1; -;  
 SQ SEQUENCE 746 AA; 83851 MW; AEC8ED567BEEB0A5 CRC64;

Query Match 46.7%; Score 7; DB 10; Length 746;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 EKLRRER 11  
 |||||  
 Db 80 EKLRRER 86

RESULT 6  
 O45358

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:35:54 ; Search time 89.06 Seconds  
(without alignments)  
22.284 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_154  
Perfect score: 15  
Sequence: 1 HSSKELRRRIKYC 15

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_16:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_unclassified:\*
- 13: sp\_vertebrate:\*
- 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	60.0	425	4 Q9NX45	Q9NX45 homo sapien
2	7	46.7	250	10 Q9MOV5	Q9MOV5 arabidopsis
3	7	46.7	443	10 Q9LYR0	Q9LYR0 arabidopsis
4	7	46.7	502	2 Q9ZB31	Q9ZB31 bruceella ab
5	7	46.7	746	10 Q9FNF6	Q9FNF6 arabidopsis
6	7	46.7	791	5 O45358	O45358 caenorhabdi
7	7	46.7	1080	11 Q9WVP5	Q9WVP5 mus musculu
8	7	46.7	1080	11 Q9EQU1	Q9EQU1 mus musculu
9	7	46.7	1260	5 Q9GRG5	Q9GRG5 trypanosoma
10	7	46.7	1341	3 Q98281	Q98281 saccharomyc
11	7	46.7	1342	3 Q92271	Q92271 saccharomyc
12	7	46.7	1833	5 Q9VMG7	Q9VMG7 drosophila
13	6	40.0	39	4 Q9UBN9	Q9UBN9 homo sapien
14	6	40.0	64	2 Q9EY25	Q9EY25 photobacter
15	6	40.0	87	2 Q9KVL1	Q9KVL1 vibrio chol
16	6	40.0	90	1 Q96538	Q96538 methanobact
17	6	40.0	94	1 Q9HH63	Q9HH63 methanobact
18	6	40.0	94	9 O80193	O80193 methanobact
19	6	40.0	131	10 Q9XES9	Q9XES9 glycine max

20	6	40.0	132	8 Q9TJN6	Q9TJN6 zea mays (m
21	6	40.0	135	5 Q17714	Q17714 caenorhabdi
22	6	40.0	160	5 Q93970	Q93970 caenorhabdi
23	6	40.0	171	10 Q9FRL4	Q9FRL4 brassica na
24	6	40.0	184	3 Q13954	Q13954 schizosacch
25	6	40.0	185	2 Q9F2S8	Q9F2S8 streptomyce
26	6	40.0	201	2 Q9RR40	Q9RR40 citrobacter
27	6	40.0	218	2 Q9RUE2	Q9RUE2 deinococcus
28	6	40.0	218	5 Q9V4H1	Q9V4H1 drosophila
29	6	40.0	219	1 Q98180	Q98180 pyrococcus
30	6	40.0	222	10 Q38918	Q38918 arabidopsis
31	6	40.0	239	2 Q9K744	Q9K744 bacillus ha
32	6	40.0	242	1 Q9UXY9	Q9UXY9 pyrococcus
33	6	40.0	249	1 Q9V0V8	Q9V0V8 pyrococcus
34	6	40.0	267	5 Q9XZ21	Q9XZ21 leishmania
35	6	40.0	267	14 Q9J8B0	Q9J8B0 spodoptera
36	6	40.0	281	5 Q45765	Q45765 caenorhabdi
37	6	40.0	282	2 Q9PN93	Q9PN93 campylobact
38	6	40.0	288	13 Q9YH08	Q9YH08 xenopus lae
39	6	40.0	290	1 Q28618	Q28618 archaetoglob
40	6	40.0	296	2 Q9I4F4	Q9I4F4 pseudomonas
41	6	40.0	297	10 Q9SA66	Q9SA66 arabidopsis
42	6	40.0	297	10 Q9LSQ3	Q9LSQ3 arabidopsis
43	6	40.0	310	10 Q22768	Q22768 arabidopsis
44	6	40.0	311	5 Q9W019	Q9W019 drosophila
45	6	40.0	314	5 Q9VQ73	Q9VQ73 drosophila

#### ALIGNMENTS

RESULT 1

Q9NX45 ID Q9NX45 PRELIMINARY; PRT; 425 AA.  
AC Q9NX45;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE CDNA FLJ20449 FIS, CLONE RAT05575.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,  
RA Suzuki Y., Ohbayashi M., Nishi T., Shibahara T., Tanaka T.,  
RA Nakamura Y., Isogai T., Sugano S.;  
RT "NEDO human cDNA sequencing project."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK000456; BAA91175.1; -.  
DR InterPro; IPR001092; -.  
DR SMART; SM00353; HLH; 1.  
SQ SEQUENCE 425 AA; 46922 MW; 777E57231F9D6BDD CRC64;

Query Match 60.0%; Score 9; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 0.37;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 LRRRIKYC 15  
| | | | | | | | | |  
Db 212 LRRRIKYC 220

RESULT 2

Q9MOV5 ID Q9MOV5 PRELIMINARY; PRT; 250 AA.  
AC Q9MOV5;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE HYPOTHETICAL 27.6 KDA PROTEIN.



**THIS PAGE BLANK (USPTO)**



PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149375.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151086.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.

PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 40.0%; Score 6; DB 21; Length 259;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRERI 12  
Db 112 lrreri 117

Search completed: September 15, 2001, 12:39:12  
Job time: 2893 sec

PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 40.0%; Score 6; DB 21; Length 259;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRERI 12  
Db 112 Lrreri 117

RESULT 15  
AAG45159  
ID AAG45159 standard; Protein; 259 AA.

XX AC AAG45159;

XX XX 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 56659.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX XX 25-FEB-2000; 2000EP-0301439.

XX XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 21-APR-1999; 99US-0130077.

PR 23-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.

PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.

```
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145216.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 40.0%; Score 6; DB 21; Length 253;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRERI 12
Db 112 lrreri 117

RESULT 14
AAG05356
ID AAG05356 standard; Protein; 259 AA.
XX
AC AAG05356;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1732.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
```

PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 14-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 22-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162143.

Query Match 40.08; Score 6; DB 21; Length 252;  
Best Local Similarity 100.08; Pred. No. 67;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERI 12  
DQ 111 Lrreri 116

RESULT 13  
AAG39231  
ID AAG39231 standard; Protein: 253 AA.

XX AC AAG39231;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 48511.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX XX 06-SEP-2000.

XX XX 25-FEB-2000; 2000EP-0301439.

XX XX

PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 21-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141267.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.

XX EP1033405-A2.  
PN 01-JUL-1999; 99US-0141842.  
XX 02-JUL-1999; 99US-0142154.  
XX 06-JUL-1999; 99US-0142055.  
PD 06-SEP-2000. 99US-0142390.  
XX 08-JUL-1999; 99US-0142803.  
PF 09-JUL-1999; 99US-0142920.  
PF 12-JUL-1999; 99US-0142977.  
XX 13-JUL-1999; 99US-0143542.  
XX 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151085.  
PR 27-AUG-1999; 99US-0151086.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.



PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157855.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161922.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 40.0%; Score 6; DB 21; Length 248;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRERI 12  
 |||||  
 Db 107 Lrrerl 112

RESULT 11  
 AAB42198  
 ID AAB42198 standard; Protein; 250 AA.

XX AC AAB42198;  
 XX DT 08-FEB-2001 (first entry)  
 XX DE Human ORFX ORF1962 polypeptide sequence SEQ ID NO:3924.  
 XX DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX OS Homo sapiens.  
 XX

PN WO200058473-A2.

XX PD 05-OCT-2000.

XX PF 31-MAR-2000; 2000WO-US08621.

XX PR 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.

PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX N-PSDB; AAC76407.

XX Novel nucleic acids and peptides derived from open reading frame X,  
 useful for treating e.g. cancers, proliferative disorders,  
 neurodegenerative disorders and cardiovascular disease -

Claim 11; Page 3086-3087; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 sequences have activities such as: cytostatic; hepatotropic; vulnary;  
 antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
 immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 antidiabetic; hypotensive; dermatological; immunosuppressive;  
 antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;  
 antithyroid; and antianaemic. The sequences can be used for determining  
 the presence of or predisposition to, or preventing or treating  
 pathological conditions associated with an ORFX-associated disorder. The  
 nucleic acids can be used to express ORFX proteins in gene therapy  
 vectors. The proteins and nucleic acids may be used to treat cancers,  
 proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 graft vs host disease, cardiovascular disease, diabetes mellitus,  
 hyperthyroidism, hypothyroidism, cholesterol ester storage, systemic lupus  
 erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
 coagulation; to inhibit thrombosis; and as a contraceptive.

SQ Sequence 250 AA;

Query Match 40.0%; Score 6; DB 21; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSSREK 6  
 |||||  
 Db 12 hsssek 17

RESULT 12  
 AAG39232  
 ID AAG39232 standard; Protein; 252 AA.

XX AC AAG39232;

XX DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 48512.

DE Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX Arabidopsis thaliana.  
 OS



```
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142300.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0150894.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 40.0%; Score 6; DB 21; Length 247;
Best Local Similarity 100.0%; Pred.No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRERI 12
Db 106 lrreri lll

RESULT 10
AAG30796
ID AAG30796 standard; Protein; 248 AA.
XX
AC AAG30796;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 36881.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
```

```
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156590.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 40.0%; Score 6; DB 21; Length 152;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SKEKLK 8
Db 98 skeklr 103

RESULT 9
AAG30797
ID AAG30797 standard; Protein; 247 AA.
XX
AC AAG30797;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 36882.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
```

```
XX Arabidopsis thaliana.
OS
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
```

DT	18-OCR-2000 (first entry)	PR	18-JUN-1999;	99US-0139763.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 70291.	PR	21-JUN-1999;	99US-0139817.
XX		PR	22-JUN-1999;	99US-0139899.
XX	Protein identification; signal transduction pathway; metabolic pathway;	PR	23-JUN-1999;	99US-0140353.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	PR	24-JUN-1999;	99US-0140354.
KW	termination sequence.	PR	28-JUN-1999;	99US-0140695.
XX		PR	29-JUN-1999;	99US-0140823.
OS	Arabidopsis thaliana.	PR	01-JUL-1999;	99US-0140991.
XX		PR	01-JUL-1999;	99US-0141287.
PN	EP1033405-A2.	PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
XX		PR	06-JUL-1999;	99US-0142350.
PD	06-SEP-2000.	PR	08-JUL-1999;	99US-0142803.
PF	25-FEB-2000; 2000EP-0301439.	PR	09-JUL-1999;	99US-0142920.
XX		PR	12-JUL-1999;	99US-0142977.
PR	25-FEB-1999;	PR	13-JUL-1999;	99US-0143542.
PR	05-MAR-1999;	PR	14-JUL-1999;	99US-0143624.
PR	09-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	25-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	19-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.

PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140333.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147433.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 46.7%; Score 7; DB 21; Length 443;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 LRERIK 13  
| | | | |  
Db 227 lrerik 233

RESULT 8  
AAG54945  
ID AAG54945 standard; Protein; 152 AA.  
XX  
AC AAG54945;  
XX

```
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 46.7%; Score 7; DB 21; Length 341;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 LRRERIK 13
Db 125 lrrerik 131
```

---

```
RESULT 7
AAG15756
ID AAG15756 standard; Protein; 443 AA.
XX
AC AAG15756;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 16133.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
```

Best Local Similarity 100.0%; Pred. No. 7.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERIK 13  
Db 75 lrrerik 81

RESULT 6  
AAG15757  
ID AAG15757 standard; Protein; 341 AA.  
XX  
AC AAG15757;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 16134.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123546.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.



PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0147038.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 04-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 26-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match

46.7%; Score 7; DB 21; Length 291;

```
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 46.7%; Score 7; DB 21; Length 250;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKELR 8
Db 244 sskelr 250
|||||

RESULT 5
AAG15758
ID AAG15758 standard; Protein; 291 AA.
XX
AC AAG15758;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 16135.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
```

```
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.

Query Match 46.7%; Score 7; DB 21; Length 135;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKKLIR 8
Db 129 sskelir 135

RESULT 4
ID AAG43790 standard; Protein; 250 AA.
XX AC
XX AAG43790;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 54774.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126284.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 04-MAY-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 06-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 11-MAY-1999; 99US-0132863.
XX PR 14-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
```

PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 06-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 11:50:59 ; Search time 50.94 Seconds  
(without alignments)  
17.852 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_154  
Perfect score: 15  
Sequence: 1 HSSKEKLRRRIKYC 15

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_0601.\*  
1: /cgnl\_9/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /cgnl\_9/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /cgnl\_9/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /cgnl\_9/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /cgnl\_9/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /cgnl\_9/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /cgnl\_9/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /cgnl\_9/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /cgnl\_9/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /cgnl\_9/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /cgnl\_9/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /cgnl\_9/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /cgnl\_9/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /cgnl\_9/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /cgnl\_9/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /cgnl\_9/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /cgnl\_9/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /cgnl\_9/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /cgnl\_9/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /cgnl\_9/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /cgnl\_9/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /cgnl\_9/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	405	21	AAV79269 Human testis-speci
2	14	93.3	15	21	AAV79270 PHELIx peptide use
3	7	46.7	135	21	AAAG43791 Arabidopsis thalia
4	7	46.7	250	21	AAAG43790 Arabidopsis thalia
5	7	46.7	291	21	AAAG15758 Arabidopsis thalia
6	7	46.7	341	21	AAAG15757 Arabidopsis thalia
7	7	46.7	443	21	AAAG15756 Arabidopsis thalia
8	6	40.0	152	21	AAAG54945 Arabidopsis thalia
9	6	40.0	247	21	AAAG30797 Arabidopsis thalia
10	6	40.0	248	21	AAAG30796 Arabidopsis thalia
11	6	40.0	250	21	AAB42198 Human ORFX ORF1962

12	6	40.0	252	21	AAG39232 Arabidopsis thalia
13	6	40.0	253	21	AAG39231 Arabidopsis thalia
14	6	40.0	259	21	AAG05356 Arabidopsis thalia
15	6	40.0	259	21	AAG45159 Arabidopsis thalia
16	6	40.0	260	21	AAG05355 Arabidopsis thalia
17	6	40.0	260	21	AAG45158 Arabidopsis thalia
18	6	40.0	284	21	AAG22161 Arabidopsis thalia
19	6	40.0	291	21	AAG18003 Arabidopsis thalia
20	6	40.0	291	21	AAG49791 Arabidopsis thalia
21	6	40.0	293	19	AAAG69419 Protein encoded by
22	6	40.0	297	21	AAAG30795 Arabidopsis thalia
23	6	40.0	302	21	AAG39230 Arabidopsis thalia
24	6	40.0	305	21	AAB56616 Human prostate can
25	6	40.0	310	21	AAG05354 Arabidopsis thalia
26	6	40.0	310	21	AAG45157 Arabidopsis thalia
27	6	40.0	311	21	AAG49790 Arabidopsis thalia
28	6	40.0	320	21	AAAG18002 Arabidopsis thalia
29	6	40.0	329	21	AAAG22160 Arabidopsis thalia
30	6	40.0	333	21	AAAG22159 Arabidopsis thalia
31	6	40.0	337	22	AAB60661 Human apoptosis-re
32	6	40.0	348	21	AAG18001 Arabidopsis thalia
33	6	40.0	360	21	AAG49789 Arabidopsis thalia
34	6	40.0	397	21	AAB26437 Drosophila melanog
35	6	40.0	397	21	AAB20934 Drosophila odorant
36	6	40.0	439	21	AAV84905 A human proliferat
37	6	40.0	481	19	AAW35000 Archaeobacterium AE
38	6	40.0	524	20	AAW97815 Guinea pig butyrop
39	6	40.0	526	20	AAW97812 Bovine butyrophill
40	6	40.0	866	17	AAW03729 Human papilloma vi
41	6	40.0	866	20	AAW08756 HPV E6-AP protein.
42	6	40.0	874	16	AAW79657 Human E6-AP protei
43	6	40.0	874	17	AAW06358 E6-associated prot
44	6	40.0	874	20	AAV39969 Human E6-AP protei
45	6	40.0	874	21	AAB03177 Human E6-associate

ALIGNMENTS

RESULT 1  
AAV79269  
ID AAV79269 standard; Protein; 405 AA.  
XX AC AAV79269;  
XX 03-JUL-2000 (first entry)  
XX Human testis-specific transcription factor PHELIx.  
DE PHELIx; human; testis-specific; transcription factor;  
KW prostate cancer; bladder cancer; ovary cancer; testicular cancer;  
KW therapy; diagnosis; vaccine.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FT Peptide 134..150  
FT Peptide /note= "nuclear localization signal"  
FT Peptide 163..169  
FT Peptide /note= "nuclear localization signal"  
FT Domain 140..189  
FT /note= "basic Helix-Loop-Helix domain"  
XX WO200012709-A2.  
PN 09-MAR-2000.  
XX 31-AUG-1999; 99WO-US20137.  
XX 31-AUG-1998; 98US-0098610.  
PR 31-OCT-1998; 98US-0106524.  
XX (UROC-) UROGENESYS INC.  
PA

**THIS PAGE BLANK (USPTO)**

ADDRESSEE: ROGERS, HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
CITY: ST. LOUIS  
STATE: MISSOURI  
COUNTRY: US  
ZIP: 63105-1817  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/742.035  
FILING DATE: 01-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/519,777  
FILING DATE: 28-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 953095  
TELEPHONE: (314) 727-5188  
TELEFAX: (314) 727-6092  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 94 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-742-035-31

Query Match 33.3%; Score 5; DB 1; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 7 LRRR 11  
Db 52 LRRR 56

Search completed: September 15, 2001, 12:40:19  
Job time: 1640 sec



; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/FR93/00494  
 ; FILING DATE: 19-MAY-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: FR 92/06123  
 ; FILING DATE: 20-MAY-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Weiser, Gerard J.  
 ; REGISTRATION NUMBER: 19,763  
 ; REFERENCE/DOCKET NUMBER: 989.6121P  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215-875-8383  
 ; TELEFAX: 215-875-8394  
 ; INFORMATION FOR SEQ ID NO: 10:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 78 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-343-443B-10

Query Match 33.3%; Score 5; DB 2; Length 78;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 KLRRE 10  
 Db 14 KLRRE 18

RESULT 13  
 US-07-667-276A-9  
 ; Sequence 9, Application US/07667276A  
 ; Patent No. 5470971  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kondo, Keiji  
 ; APPLICANT: Inouye, Masayori  
 ; TITLE OF INVENTION: STRESS-INDUCED PROTEINS, GENES CODING  
 ; THEREFOR, TRANSFORMED CELLS OF ORGANISMS, METHODS AND  
 ; TITLE OF INVENTION: APPLICATIONS  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Weiser & Associates  
 ; STREET: 230 S. Fifteenth Street, Suite 500  
 ; CITY: Philadelphia  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19102  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07667,276A  
 ; FILING DATE: 11-MAR-1991  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Weiser, Gerard J.  
 ; REGISTRATION NUMBER: 19,763  
 ; REFERENCE/DOCKET NUMBER: 377.5351P  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 215-875-8383  
 ; TELEFAX: 215-875-8394  
 ; INFORMATION FOR SEQ ID NO: 9:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 85 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; US-07-667-276A-9

Query Match 33.3%; Score 5; DB 1; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 6 KLRRE 10  
 Db 17 KLRRE 21

RESULT 14  
 US-08-519-777-31  
 ; Sequence 31, Application US/08519777  
 ; Patent No. 5739307  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHNSON JR., EUGENE M.  
 ; APPLICANT: MILBRANDT, JEFFREY D.  
 ; APPLICANT: KOTZBAUER, PAUL T.  
 ; APPLICANT: LAMPE, PATRICIA A.  
 ; TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS  
 ; NUMBER OF SEQUENCES: 78  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: ROGERS, HOWELL & HAFERKAMP, L.C.  
 ; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400  
 ; CITY: ST. LOUIS  
 ; STATE: MISSOURI  
 ; COUNTRY: US  
 ; ZIP: 63105-1817  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/519,777  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: HOLLAND, DONALD R.  
 ; REGISTRATION NUMBER: 35,197  
 ; REFERENCE/DOCKET NUMBER: 953095  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (314) 727-5188  
 ; TELEFAX: (314) 727-6092  
 ; INFORMATION FOR SEQ ID NO: 31:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 94 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-519-777-31

Query Match 33.3%; Score 5; DB 1; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRER 11  
 Db 52 LRRER 56

RESULT 15  
 US-08-742-035-31  
 ; Sequence 31, Application US/08742035  
 ; Patent No. 5747655  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHNSON JR., EUGENE M.  
 ; APPLICANT: MILBRANDT, JEFFREY D.  
 ; APPLICANT: KOTZBAUER, PAUL T.  
 ; APPLICANT: LAMPE, PATRICIA A.  
 ; TITLE OF INVENTION: NEURTURIN AND RELATED GROWTH FACTORS  
 ; NUMBER OF SEQUENCES: 78  
 ; CORRESPONDENCE ADDRESS:

;; APPLICANT: Chung, Soo-Il  
;; APPLICANT: Park, Sang-Chul  
;; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and  
;; TITLE OF INVENTION: Methods of Using Same  
;; NUMBER OF SEQUENCES: 117  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Knobbe, Martens, Olson & Bear  
;; STREET: 620 Newport Center Drive, Sixteenth Floor  
;; CITY: Newport Beach  
;; STATE: CA  
;; COUNTRY: U.S.A.  
;; ZIP: 92660  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/056,200  
;; FILING DATE: 30-APR-1993  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fedrick, Michael F.  
;; REGISTRATION NUMBER: 36,799  
;; REFERENCE/DOCKET NUMBER: NIH054.001A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (714) 760-0404  
;; TELEFAX: (714) 760-9502  
;; INFORMATION FOR SEQ ID NO: 15:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 26 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;; US-08-056-200-15

Query Match 33.3%; Score 5; DB 1; Length 26;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 RERIK 13  
|||||  
DB 5 RERIK 9

RESULT 11  
US-08-644-15  
; Sequence 15, Application US/08800644  
; Patent No. 5958752  
; GENERAL INFORMATION:  
; APPLICANT: Steinert, Peter M.  
; APPLICANT: Lee, Seung-Chul  
; APPLICANT: Kim, In-Gyu  
; APPLICANT: Chung, Soo-Il  
; APPLICANT: Park, Sang-Chul  
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and  
; TITLE OF INVENTION: Methods of Using Same  
; NUMBER OF SEQUENCES: 117  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/800,644  
;; FILING DATE: 14-FEB-1997  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/056,200  
;; FILING DATE: 30-APR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fedrick, Michael F.  
;; REGISTRATION NUMBER: 36,799  
;; REFERENCE/DOCKET NUMBER: NIH054.001A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (714) 760-0404  
;; TELEFAX: (714) 760-9502  
;; INFORMATION FOR SEQ ID NO: 15:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 26 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
;; US-08-800-644-15

Query Match 33.3%; Score 5; DB 2; Length 26;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 RERIK 13  
|||||  
DB 5 RERIK 9

RESULT 12  
US-08-343-443B-10  
; Sequence 10, Application US/08343443B  
; Patent No. 5968734  
; GENERAL INFORMATION:  
; APPLICANT: Aurias, Alain  
; APPLICANT: Delattre, Olivier  
; APPLICANT: Desmaze, Chantal  
; APPLICANT: Melot, Thomas  
; APPLICANT: Peter, Martine  
; APPLICANT: Ploougastel, Beatrice  
; APPLICANT: Thomas, Gilles  
; APPLICANT: Zucman, Jessica  
; TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF  
; TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL  
; TITLE OF INVENTION: TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS  
; TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID  
; TITLE OF INVENTION: TRANSLOCATIONS  
; NUMBER OF SEQUENCES: 129  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weiser & Associates  
; STREET: 230 South Fifteenth Street  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: AEDIT 1.0 DOS text editor  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/343,443B  
; FILING DATE: 18-NOV-1994  
; CLASSIFICATION: 514

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US93/12679  
;; FILING DATE: 30-DEC-1993  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Cranfill, Raymond B  
;; REGISTRATION NUMBER: 32,845  
;; REFERENCE/DOCKET NUMBER: RATH-10016PC  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-322-5333  
;; TELEFAX: 415-322-5499  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 6 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; PCT-US93-12679-5

Query Match 33.3%; Score 5; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSKEK 6  
Db 2 SSKEK 6

;; RESULT 8  
;; US-08-704-170-28  
;; Sequence 28, Application US/08704170  
;; Patent No. 5707626  
;; GENERAL INFORMATION:  
;; APPLICANT: Douvas, Angeline  
;; APPLICANT: Takehana, Yoshi  
;; APPLICANT: Ehresmann, Glenn  
;; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
;; NUMBER OF SEQUENCES: 121  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Robbins, Berliner & Carson  
;; STREET: 201 No. 5707626th Figueroa Street, Suite 500  
;; CITY: Los Angeles  
;; STATE: California  
;; COUNTRY: U.S.A.  
;; ZIP: 90012  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING-SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/704,170  
;; FILING DATE:  
;; CLASSIFICATION: 424

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/029,850  
;; FILING DATE: 11-MAR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Spitals, John P.  
;; REGISTRATION NUMBER: 29,215  
;; REFERENCE/DOCKET NUMBER: 1920-331  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (213) 977-1001  
;; TELEFAX: (213) 977-1003  
;; INFORMATION FOR SEQ ID NO: 28:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 21 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-704-170-28

Query Match 33.3%; Score 5; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSKEK 6

;; Query Match 33.3%; Score 5; DB 1; Length 21;  
;; Best Local Similarity 100.0%; Pred. No. 32;  
;; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
;; Qy 6 KLRRE 10  
;; Db 2 KLRRE 6  
;; RESULT 9  
;; PCT-US94-02631-28  
;; Sequence 28, Application PC/TUS9402631  
;; GENERAL INFORMATION:  
;; APPLICANT: Douvas, Angeline  
;; APPLICANT: Takehana, Yoshi  
;; APPLICANT: Ehresmann, Glenn  
;; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
;; NUMBER OF SEQUENCES: 121  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Robbins, Berliner & Carson  
;; STREET: 201 North Figueroa Street, Suite 500  
;; CITY: Los Angeles  
;; STATE: California  
;; COUNTRY: U.S.A.  
;; ZIP: 90012  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US94/02631  
;; FILING DATE:  
;; CLASSIFICATION:

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/029,850  
;; FILING DATE: 11-MAR-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Spitals, John P.  
;; REGISTRATION NUMBER: 29,215  
;; REFERENCE/DOCKET NUMBER: 1920-331  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (213) 977-1001  
;; TELEFAX: (213) 977-1003  
;; INFORMATION FOR SEQ ID NO: 28:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 21 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; PCT-US94-02631-28

Query Match 33.3%; Score 5; DB 5; Length 21;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 KLRRE 10  
Db 2 KLRRE 6

;; RESULT 10  
;; US-08-056-200-15  
;; Sequence 15, Application US/08056200  
;; Patent No. 5616500  
;; GENERAL INFORMATION:  
;; APPLICANT: Steinert, Peter M.  
;; APPLICANT: Lee, Seung-Chul  
;; APPLICANT: Kim, In-Gyu

; TELEPHONE: (206) 467-9600  
 ; TELEFAX: (415) 543-5043  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 866 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-674-030-1

Query Match 40.0%; Score 6; DB 2; Length 866;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSKEKL 7  
 Db 844 SSKEKL 849

## RESULT 5

US-08-247-904B-8  
 ; Sequence 8, Application US/08247904B  
 ; Patent No. 5981699

; GENERAL INFORMATION:  
 ; APPLICANT: Rolfe, Mark  
 ; APPLICANT: Eckstein, Jens W.  
 ; APPLICANT: Draetta, Giulio  
 ; TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme  
 ; NUMBER OF SEQUENCES: 17  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley, Hoag & Eliot  
 ; STREET: One Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII(text)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/247,904B  
 ; FILING DATE: 23-MAY-1994  
 ; CLASSIFICATION: 530  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Vincent, Matthew P.  
 ; REGISTRATION NUMBER: 36,709  
 ; REFERENCE/DOCKET NUMBER: MIV-029.01

; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 832-1000  
 ; TELEFAX: (617) 832-7000  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 874 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-247-904B-8

Query Match 40.0%; Score 6; DB 2; Length 874;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSKEKL 7  
 Db 852 SSKEKL 857

## RESULT 6

US-08-767-942A-21  
 ; Sequence 21, Application US/08767942A  
 ; Patent No. 6068982  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rolfe, Mark  
 ; APPLICANT: Chiu, M. Isabel  
 ; APPLICANT: Berlin, Vivian  
 ; APPLICANT: Damagnez, Veronique  
 ; APPLICANT: Draetta, Giulio  
 ; APPLICANT: Guillaume, Cottarel  
 ; TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES  
 ; NUMBER OF SEQUENCES: 45  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
 ; STREET: One Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109-2170

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/767,942A  
 ; FILING DATE: 17-DEC-1996

; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Vincent, Matthew P.  
 ; REGISTRATION NUMBER: 36,709  
 ; REFERENCE/DOCKET NUMBER: MIV-029.04  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-832-1000  
 ; TELEFAX: 617-832-7000  
 ; INFORMATION FOR SEQ ID NO: 21:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 874 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-767-942A-21

Query Match 40.0%; Score 6; DB 3; Length 874;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSKEKL 7  
 Db 852 SSKEKL 857

## RESULT 7

PCT-US93-12679-5

; Sequence 5, Application PC/TUS9312679  
 ; GENERAL INFORMATION:

; APPLICANT: Rath, Matthias  
 ; TITLE OF INVENTION: SYNTHETIC OLIGOPEPTIDES ANALOGOUS TO  
 ; TITLE OF INVENTION: PROTEIN SIGNAL SEQUENCES -METHODS OF USE  
 ; NUMBER OF SEQUENCES: 41  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SHELDON & MAK  
 ; STREET: 401 Florence Street  
 ; CITY: Palo Alto  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94301

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25

```
Db      30 SSKEKLR 36

RESULT 2
US-08-611-510-6
; Sequence 6, Application US/08611510
; Patent No. 5824529
; GENERAL INFORMATION:
; APPLICANT: Morgan, Zhiyuh
; APPLICANT: Chang, Richard D.
; TITLE OF INVENTION: Method For Cloning And
; TITLE OF INVENTION: Producing The Pshai Restriction
; TITLE OF INVENTION: Endonuclease
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gregory D. Williams; NEW ENGLAND
; ADDRESSEE: BIOLABS, INC.
; STREET: 32 Tozer Road
; CITY: Beverly
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 01915
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,510
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: NEB-119
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 927-5054
; TELEFAX: (508) 927-1705
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS: not applicable
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-611-510-6

Query Match      40.0%; Score 6; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      7 LRRERI 12
        |||||
Db      16 LRRERI 21

RESULT 3
US-08-100-692-1
; Sequence 1, Application US/08100692
; Patent No. 5532348
; GENERAL INFORMATION:
; APPLICANT: Hubbrechtse, Jon M.
; APPLICANT: Scheffner, Martin
; APPLICANT: Howley, Peter M.
; TITLE OF INVENTION: B6 ASSOCIATED PROTEIN AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco

Query Match      40.0%; Score 6; DB 1; Length 866;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 SSKEKL 7
        |||||
Db      844 SSKEKL 849

RESULT 4
US-08-674-030-1
; Sequence 1, Application US/08674030
; Patent No. 5914389
; GENERAL INFORMATION:
; APPLICANT: Hubbrechtse, Jon M.
; APPLICANT: Scheffner, Martin
; APPLICANT: Howley, Peter M.
; TITLE OF INVENTION: B6 ASSOCIATED PROTEIN AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/674,030
; FILING DATE: 01-JUL-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,692
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 15280-91
; TELECOMMUNICATION INFORMATION:
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:12:59 ; Search time 57.85 Seconds  
(without alignments)  
5.339 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_154  
Perfect score: 15  
Sequence: 1 HSSKEKLRRRIKYC 15

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgnl\_7/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgnl\_7/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgnl\_7/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgnl\_7/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgnl\_7/ptodata/1/iaa/PCTUS.COMB.pep.\*  
6: /cgnl\_7/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	46.7	52	3	US-08-630-916A-76
2	6	40.0	79	2	US-08-611-510-6
3	6	40.0	866	1	US-08-100-692-1
4	6	40.0	866	2	US-08-674-030-1
5	6	40.0	874	2	US-08-247-904B-8
6	6	40.0	874	3	US-08-767-942A-21
7	5	33.3	6	5	PCT-US93-12679-5
8	5	33.3	21	1	US-08-704-170-28
9	5	33.3	26	1	PCT-US94-02631-28
10	5	33.3	26	2	US-08-056-200-15
11	5	33.3	26	2	US-08-800-644-15
12	5	33.3	78	2	US-08-343-443B-10
13	5	33.3	85	1	US-07-667-276A-9
14	5	33.3	94	1	US-08-519-777-31
15	5	33.3	94	1	US-08-742-035-31
16	5	33.3	94	2	US-08-777-019-31
17	5	33.3	94	2	US-08-777-143-31
18	5	33.3	94	3	US-08-775-414-31
19	5	33.3	94	4	US-08-931-858E-31
20	5	33.3	94	4	US-08-981-739-31
21	5	33.3	95	4	US-08-931-858E-173
22	5	33.3	95	4	US-08-981-739-173
23	5	33.3	102	1	US-08-519-777-1
24	5	33.3	102	1	US-08-742-035-1
25	5	33.3	102	2	US-08-777-019-1
26	5	33.3	102	2	US-08-777-143-1
27	5	33.3	102	3	US-08-478-097A-18

28	5	33.3	102	3	US-09-106-486-1	Sequence 1, Appli
29	5	33.3	102	3	US-08-775-414-1	Sequence 1, Appli
30	5	33.3	102	4	US-08-931-858E-1	Sequence 1, Appli
31	5	33.3	102	4	US-08-981-739-1	Sequence 1, Appli
32	5	33.3	103	3	US-09-106-486-4	Sequence 4, Appli
33	5	33.3	104	3	US-08-775-414-79	Sequence 79, Appli
34	5	33.3	129	6	5196523-10	Patent No. 5196523
35	5	33.3	144	3	US-08-775-414-81	Sequence 81, Appli
36	5	33.3	148	5	PCT-US95-07135-2	Sequence 2, Appli
37	5	33.3	152	3	US-08-775-414-83	Sequence 83, Appli
38	5	33.3	173	3	US-08-937-271-15	Sequence 15, Appli
39	5	33.3	197	1	US-08-519-777-7	Sequence 7, Appli
40	5	33.3	197	1	US-08-742-035-7	Sequence 7, Appli
41	5	33.3	197	2	US-08-777-019-7	Sequence 7, Appli
42	5	33.3	197	2	US-08-777-143-7	Sequence 7, Appli
43	5	33.3	197	3	US-09-106-486-5	Sequence 5, Appli
44	5	33.3	197	3	US-08-775-414-7	Sequence 7, Appli
45	5	33.3	197	4	US-08-931-858E-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1  
US-08-630-916A-76  
; Sequence 76, Application US/08630916A  
; Patent No. 6011137  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; APPLICANT: Kay, Brian K.  
; APPLICANT: Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/630,916A  
; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 896-8864/9741  
; INFORMATION FOR SEQ ID NO: 76:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 52 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-630-916A-76

Query Match 46.7%; Score 7; DB 3; Length 52;  
Best Local Similarity 100.0%; Pred. No. 0.68;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 SSKKLRR 8  
|||||

**THIS PAGE BLANK (USPTO)**

A:Reference number: Z19877  
A:Accession: T24339  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-160 <WIL>  
A:Cross-references: EMBL:Z80219; PIDN:CA802300.1; GSPDB:GN00019; CESP:T01H8.3  
A:Experimental source: clone T01H8  
C:Genetics:  
A:Gene: CESP:T01H8.3  
A:Map position: 1  
A:Introns: 40/3; 79/1; 128/3

Query Match 40.0%; Score 6; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 5 EKLRR 10  
    |||||  
Db 6 EKLRR 11

Search completed: September 15, 2001, 12:41:25  
Job time: 379 sec



A:Reference number: A69000; MUID:98037514  
A:Accession: H69159  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-90 <MTH>  
A:Cross-references: GB:AE000830; GB:AE000666; NID:g2621523; PIDN:AAB84964.1; PID:g262152  
A:Experimental source: strain Delta H  
A:Genetics:  
A:Gene: MTH458

Query Match 40.0%; Score 6; DB 2; Length 90;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KEKLR 9  
|||||

Db 36 KEKLR 41

RESULT 11  
D84066  
host factor I - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 29-Sep-1999  
C:Accession: D84066  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A. Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: D84066  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-91 <TIGR>  
A:Cross-references: GB:U32724; GB:L42023; NID:g1573378; PIDN:AAC22070.1; PID:g1573384; T  
C:Superfamily: host factor I

Query Match 40.0%; Score 6; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERI 12  
|||||

Db 15 LRRERI 20

RESULT 12  
T12719  
hypothetical protein 3 - Methanobacterium phage psiM2  
C:Species: Methanobacterium phage psiM2  
C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 24-Sep-1999  
C:Accession: T12719  
R:Pfister, P.; Wasserfallen, A.; Stettler, R.; Leisinger, T.  
submitted to the EMBL Data Library, May 1998  
A:Description: Archaeophage PsiM2 complete genomic DNA.  
A:Reference number: Z17578  
A:Accession: T12719  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-94 <PFI>  
A:Cross-references: EMBL:AF065411; NID:g3249585; PID:g3249588; PIDN:AAC27042.1  
A:Experimental source: host Methanobacterium thermoautotrophicum strain Marburg

Query Match 40.0%; Score 6; DB 2; Length 94;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SKEKLR 8

Db 72 SKEKLR 77

RESULT 13  
T19002  
hypothetical protein C06C3.7 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T19002  
R:Berks, M.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: Z19058  
A:Accession: T19002  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-135 <WIL>  
A:Cross-references: EMBL:Z36719; PIDN:CAA85314.1; GSPDB:GNO0020; CESP:C06C3.7  
A:Experimental source: clone C06C3  
C:Genetics:  
A:Gene: CESP:C06C3.7  
A:Map position: 2  
A:Introns: 20/1; 72/3

Query Match 40.0%; Score 6; DB 2; Length 135;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EKLRR 10  
|||||

Db 87 EKLRR 92

RESULT 14  
G84974  
transcription elongation factor greA [imported] - Buchnera sp. (strain APS)  
C:Species: Buchnera sp.  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
C:Accession: G84974  
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
Nature 407, 81-86, 2000  
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp.  
A:Reference number: A84930; MUID:20445173  
A:Accession: G84974  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-159 <STO>  
A:Cross-references: GB:AF000398; GSPDB:GNO0144  
A:Experimental source: strain APS  
C:Genetics:  
A:Gene: greA; BU384  
C:Superfamily: transcription elongation factor greb  
C:Keywords: transcription factor

Query Match 40.0%; Score 6; DB 2; Length 159;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EKLRR 10  
|||||

Db 13 EKLRR 18

RESULT 15  
T24339  
hypothetical protein T01H8.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T24339  
R:Lennard, N.  
submitted to the EMBL Data Library, September 1996

Best Local Similarity 100.0%; Pred. No. 12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EKLRRR 11  
Db 153 EKLRRR 159  
|||||

RESULT 7  
S66835  
probable membrane protein YOL138c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein AOD1341; hypothetical protein O0483  
C:Species: Saccharomyces cerevisiae  
C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 21-Jul-2000  
C:Accession: S66835; S72030  
R:Arino, J.; Casamayor, A.; Gamo, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Casas, C.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S66814  
A:Accession: S66835  
A:Molecule type: DNA  
A:Residues: 1-1341 <ARI>  
A:Cross-references: EMBL:Z74880; NID:gl420031; PID:e251919; PID:gl420032; MIPS:YOL138c  
A:Experimental source: strain S288C  
R:Aldea, M.; Piedrafito, L.; Casas, C.; Casamayor, A.; Khalid, H.; Balcells, L.; Arino, J.  
Yeast 12, 1053-1058, 1996  
A:Title: Sequence analysis of a 12 801 bp fragment of the left arm of yeast chromosome X  
protein and six other open reading frames.  
A:Reference number: S72030; MUID:97051593  
A:Accession: S72030  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1223 'V', 1224-1341 <ALD>  
A:Cross-references: EMBL:X95465; NID:gl628437; PID:CA64732.1; PID:gl628438  
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1996  
C:Genetics:  
A:Map position: 15L  
A:Note: YOL138c  
A:Superfamily: unassigned WD repeat proteins; WD repeat homology  
C:Keywords: transmembrane protein  
F:254-288/Domain: WD repeat homology <WD1>  
F:365-398/Domain: WD repeat homology <WD2>  
F:1178-1194/Domain: transmembrane #status predicted <TM>

Query Match 46.7%; Score 7; DB 2; Length 1341;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKEKL 8  
Db 1257 SSKEKL 1263  
|||||

RESULT 8  
IABY3  
protease A inhibitor 3 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein Y8010.04c; protein YMR174c  
C:Species: Saccharomyces cerevisiae  
C>Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 21-Jul-2000  
C:Accession: A01334; S16692; S55121  
R:Bielermann, K.; Montali, U.; Martin, B.; Svendsen, I.; Ottesen, M.  
Carlsberg Res. Commun. 45, 225-235, 1980  
A:Title: The amino acid sequence of proteinase A inhibitor 3 from baker's yeast.  
A:Reference number: A01334  
A:Accession: A01334  
A:Molecule type: protein  
A:Residues: 1-68 <BIE>  
A:Note: nearly all of the inhibitory activity is present in the peptide consisting of re  
R:Schu, P.; Wolf, D.H.  
FEBS Lett. 283, 78-84, 1991  
A:Title: The proteinase yscA-inhibitor, I(A)(3), gene. Studies of cytoplasmic proteinase  
A:Reference number: S16692; MUID:91243884  
A:Accession: S16692

A:Molecule type: DNA  
A:Residues: 1-68 <SCH>  
A:Cross-references: EMBL:X60050; NID:g4094; PIDN:CAA42650.1; PID:g4095  
R:Churcher, C.M.  
submitted to the EMBL Data Library, June 1995  
A:Reference number: S55118  
A:Accession: S55121  
A:Molecule type: DNA  
A:Residues: 1-68 <CHU>  
A:Cross-references: EMBL:Z49808; NID:g854440; PIDN:CAA89907.1; PID:g854444; GSPDB:GNO  
A:Experimental source: strain AB972  
C:Genetics:  
A:Gene: SGD:PAI3; MIPS:YMR174c  
A:Cross-references: SGD:S0004786; MIPS:YMR174c  
A:Map position: 13R  
C:Superfamily: proteinase A inhibitor 3  
C:Keywords: acetylated amino end; blocked amino end; proteinase inhibitor  
F:1/Modified site: acetylated amino end (Met) #status experimental

Query Match 40.0%; Score 6; DB 1; Length 68;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKEKL 7  
Db 14 SSKEKL 19  
|||||

RESULT 9  
CB2334  
probable host factor-I VC0347 [imported] - Vibrio cholerae (strain N16961 serogroup O  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: CB2334  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers  
J. R.R.; Metcalnes, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: AB2035; MUID:20406833  
A:Accession: CB2334  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-87 <HEI>  
A:Cross-references: GB:AE003852; GB:AE004123; GB:AE054756; PIDN:AAF93520.1; GSPDB:GN  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC0347  
A:Map position: 1  
C:Superfamily: host factor I

Query Match 40.0%; Score 6; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERI 12  
Db 15 LRRERI 20  
|||||

RESULT 10  
H69159  
hypothetical protein MTH458 - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: H69159  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.  
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,  
K. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu

C:Superfamily: inner membrane protein malK; ATP-binding cassette homology  
 C:Keywords: ATP; binding protein-dependent transport system; membrane protein; nucleotide  
 F:37-237/Domain: ATP-binding cassette homology <ABC>  
 F:54-62/Region: nucleotide-binding motif A (P-loop)  
 F:181-185/Region: nucleotide-binding motif B

Query Match 46.7%; Score 7; DB 2; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 6.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERIK 13  
 DB 271 LRRERIK 277

## RESULT 3

T48593  
 hypothetical protein T22N19.110 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: T48593  
 R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z24490  
 A:Accession: T48593  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-443 <BEV>  
 A:Cross-references: EMBL:AL163572  
 A:Experimental source: cultivar Columbia; BAC clone T22N19  
 C:Genetics:

A:Map position: 5  
 A:Introns: 23/3; 125/3; 196/3; 240/3  
 A>Note: T22N19.110

Query Match 46.7%; Score 7; DB 2; Length 443;  
 Best Local Similarity 100.0%; Pred. No. 7.7;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRRERIK 13  
 DB 227 LRRERIK 233

## RESULT 4

F64204  
 spermidine/putrescine transport ATP-binding protein potA homolog - Mycoplasma genitalium  
 C:Species: Mycoplasma genitalium  
 C>Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 17-Mar-2000  
 C:Accession: F64204  
 R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;  
 M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.  
 C.A.; Venter, J.C.  
 Science 270, 397-403, 1995  
 A:Title: The minimal gene complement of Mycoplasma genitalium.  
 A:Reference number: A64200; MUID:96026346  
 A:Accession: F64204  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA

A:Residues: 1-486 <TTGR>  
 A:Cross-references: GB:U39683; GB:I43967; NID:g1045711; PID:g1045714; TTGR:MG042  
 A:Experimental source: strain G-37  
 C:Genetics:

A:Genetic code: SGC3  
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
 C:Keywords: ATP  
 F:1-351/Domain: ATP-binding cassette homology #status atypical <ABC>

Query Match 46.7%; Score 7; DB 2; Length 486;  
 Best Local Similarity 100.0%; Pred. No. 8.3;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 KEKLRR 10  
 DB 100 KEKLRR 106

## RESULT 5

A44337  
 kinesin-related protein KLPA - Emericella nidulans  
 N:Alternate names: kinesin-like protein, KAR3-related; KLPA protein  
 C:Species: Emericella nidulans, Aspergillus nidulans  
 C>Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 19-Jan-2001  
 C:Accession: A44337; S24830  
 R:O'Connell, M.J.; Meluh, P.B.; Rose, M.D.; Morris, N.R.  
 J. Cell Biol. 120, 153-162, 1993  
 A:Title: Suppression of the bimC4 mitotic spindle defect by deletion of klpa, a gene  
 A:Reference number: A44337; MUID:93107178  
 A:Accession: A44337  
 A:Molecule type: mRNA  
 A:Residues: 1-770 <OIC>  
 A:Cross-references: GB:X64603; NID:g2703; PIDN:CAAM5887.1; PID:g2704  
 A>Note: sequence extracted from NCBI backbone (NCBIP:121121)  
 C:Genetics:

A:Gene: klpa  
 A:Map position: 1  
 C:Superfamily: kinesin-related protein KLPA; kinesin motor domain homology  
 C:Keywords: ATP; coiled coil; microtubule binding; nucleotide binding; P-loop  
 F:422-762/Domain: kinesin motor domain homology <KMOT>  
 F:514-521/Region: nucleotide-binding motif A (P-loop)  
 F:520/Binding site: ATP (Lys) #status predicted

Query Match 46.7%; Score 7; DB 1; Length 770;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KEKLRR 10  
 DB 398 KEKLRR 404

## RESULT 6

T20815  
 hypothetical protein F26D2.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T20815; T21409  
 R:McMurray, A.  
 submitted to the EMBL Data Library, March 1997  
 A:Reference number: Z19327  
 A:Accession: T20815  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-791 <WIL>  
 A:Cross-references: EMBL:Z93377; PIDN:CA807581.1; GSPDB:GN00023; CESP:F26D2.2  
 A:Experimental source: clone F13A7  
 R:McMurray, A.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z19418  
 A:Accession: T21409  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-791 <W12>

A:Cross-references: EMBL:Z81513; PIDN:CA804186.1; GSPDB:GN00023; CESP:F26D2.2  
 A:Experimental source: clone F26D2  
 C:Genetics:

A:Gene: CESP:F26D2.2  
 A:Map position: 5  
 A:Introns: 28/3; 194/2; 195/3; 299/2; 404/1; 462/3; 500/1; 670/3; 698/3

Query Match 46.7%; Score 7; DB 2; Length 791;

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:35:06 ; Search time 52.12 Seconds  
(without alignments)  
21.923 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_154

Perfect score: 15

Sequence: 1 HSSKEKLRRRIKYC 15

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_68.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	46.7	250	H85067	hypothetical prote
2	7	46.7	327	E84188	dipeptide transpor
3	7	46.7	443	T48593	hypothetical prote
4	7	46.7	486	F64204	spermidine/putresc
5	7	46.7	770	A44337	kinesin-related pr
6	7	46.7	791	T20815	hypothetical prote
7	7	46.7	1341	S66835	probable membrane
8	6	40.0	68	IABY3	proteinase A inhib
9	6	40.0	87	C82334	probable host fact
10	6	40.0	90	H69159	hypothetical prote
11	6	40.0	91	D64066	host factor I - Ha
12	6	40.0	94	T12719	hypothetical prote
13	6	40.0	135	T19002	hypothetical prote
14	6	40.0	159	G84974	transcription elon
15	6	40.0	160	T24339	hypothetical prote
16	6	40.0	162	S52889	probable membrane
17	6	40.0	177	E71138	probable phosphori
18	6	40.0	181	JVBPNL	DNA-packaging prot
19	6	40.0	181	F64788	DNA packaging prot
20	6	40.0	181	G85689	hypothetical prote
21	6	40.0	184	T38315	very hypothetical
22	6	40.0	189	G64909	DNA packaging prot
23	6	40.0	218	C75394	ATP phosphoribosyl
24	6	40.0	219	F71155	hypothetical prote
25	6	40.0	222	S71231	geranylgeryl pyr
26	6	40.0	232	A84091	hypothetical prote
27	6	40.0	242	A75023	hypothetical prote
28	6	40.0	242	S39642	motility protein h
29	6	40.0	249	B75110	lipote-protein li

30	6	40.0	265	2	S34668	modulation protein
31	6	40.0	281	2	T24638	hypothetical prote
32	6	40.0	282	2	D81326	5,10-methylenetet
33	6	40.0	290	2	F69456	signal sequence pe
34	6	40.0	296	2	E83498	probable transcrip
35	6	40.0	297	2	B86161	hypothetical prote
36	6	40.0	310	2	T01090	hypothetical prote
37	6	40.0	320	1	B69272	conserved hypothet
38	6	40.0	325	2	B96511	probable S-ribonuc
39	6	40.0	328	2	S35336	transcription fact
40	6	40.0	334	2	S47762	dipeptide transpor
41	6	40.0	334	2	H86027	hypothetical prote
42	6	40.0	350	2	E84634	hypothetical prote
43	6	40.0	367	2	G85362	hypothetical prote
44	6	40.0	372	2	A84462	8-amino-7-oxononan
45	6	40.0	377	2	B55745	transcription fact

#### ALIGNMENTS

RESULT 1

H85067

hypothetical protein AT4G05400 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001

C:Accession: H85067

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: H85001; MUID:20083488

A:Accession: H85067

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-250 <STO>

A:Cross-references: GB:NC\_001268; MID:g7267300; PIDN:CAB81082.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4G05400

A:Map position: 4

C:Superfamily: Arabidopsis thaliana hypothetical protein F7J7.80

Query Match 46.7%; Score 7; DB 2; Length 250;

Best Local Similarity 100.0%; Pred. No. 4.9;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSKELR 8

Db 244 SSKELR 250

RESULT 2

E64188

dipeptide transport ATP-binding protein dppF - Haemophilus influenzae (strain Rd KW20

C:Species: Haemophilus influenzae

C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 20-Aug-1999

C:Accession: E64188

R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage

, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente

A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630

A:Accession: E64188

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-327 <TIGR>

A:Cross-references: GB:U32798; GB:L42023; NID:gl574110; PIDN:AAC22837.1; PID:gl574111

C:Genetics:

A:Gene: dppF

C:Function:

A:Description: probably responsible for energy-coupling to the transport system

**THIS PAGE BLANK (USPTO)**

Query Match 26.7%; Score 4; DB 6; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSK 5  
|  
|  
|  
|  
Db 2 SSK 5

## RESULT 15

US-08-342-930-5  
; Sequence 5, Application US/08342930  
; Patent No. 5821084  
; GENERAL INFORMATION:  
; APPLICANT: OLMSTED, ELIZABETH A.  
; APPLICANT: MAURO, LAURA J.  
; APPLICANT: DAVIS, ALAN R.  
; APPLICANT: DIXON, JACK E.  
; TITLE OF INVENTION: OSTEOLAST-TESTICULAR PROTEIN TYROSINE  
; TITLE OF INVENTION: PHOSPHATASE  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/342,930  
; FILING DATE: 21-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KOSKI, ANTOINETTE F.  
; REGISTRATION NUMBER: 34,202  
; REFERENCE/DOCKET NUMBER: 20344-20975.00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 813-5600  
; TELEFAX: (415) 494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-342-930-5

Query Match 26.7%; Score 4; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSK 4  
|  
|  
|  
|  
Db 2 HSK 5

Search completed: September 15, 2001, 12:45:09  
Job time: 219 sec

REGISTRATION NUMBER: 34,380  
REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 345-9100  
TELEFAX: (617) 345-9111  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-726-306A-1

Query Match 26.7%; Score 4; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSKE 5  
DB 5 SSKE 8

RESULT 12  
US-09-177-249-235  
Sequence 235, Application US/09177249  
Patent No. 6229064  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Ohad, Nir  
APPLICANT: Kiyosue, Tomohiro  
APPLICANT: Yadegari, Ramin  
APPLICANT: Margossian, Linda  
APPLICANT: Harada, John  
APPLICANT: Goldberg, Robert B.  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: Nucleic Acids that Control Seed and Fruit  
FILE REFERENCE: Development in Plants  
CURRENT APPLICATION NUMBER: US/09/177,249  
CURRENT FILING DATE: 1998-10-22  
EARLIER APPLICATION NUMBER: US 09/071,838  
EARLIER FILING DATE: 1998-05-01  
NUMBER OF SEQ ID NOS: 324  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 235  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Arabidopsis sp.  
US-09-177-249-235

Query Match 26.7%; Score 4; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 KLRR 9  
DB 1 KLRR 4

RESULT 13  
PCT-US91-08328-24  
Sequence 24, Application PC/TUS9108328  
GENERAL INFORMATION:  
APPLICANT: Ruggeri, Zaverio M.  
APPLICANT: Houghten, Richard A.  
TITLE OF INVENTION: PEPTIDES THAT INHIBIT PLATELET BINDING  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eugene Moroz, MORGAN & FINNEGAN

STREET: 345 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/08328  
FILING DATE: 19911107  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/610,363  
FILING DATE: 07-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Moroz, Eugene  
REGISTRATION NUMBER: 25,237  
REFERENCE/DOCKET NUMBER: 1198 4079PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)758-4800  
TELEFAX: (212)751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: Cross-links  
LOCATION: 4...>10  
OTHER INFORMATION: /note="Sequence linked by  
OTHER INFORMATION: Interchain amide bond at Glu residue with Lys  
OTHER INFORMATION: residue on Arg3-Lys-Arg-Ser-Arg-Gly-Asp-Val"  
PUBLICATION INFORMATION:  
DOCUMENT NUMBER: US 4,683,291  
FILING DATE: 28-OCT-1985  
PUBLICATION DATE: 28-JUL-1987  
PUBLICATION INFORMATION:  
DOCUMENT NUMBER: US B1 4,683,291  
FILING DATE: 28-OCT-1985  
PUBLICATION DATE: 03-JUL-1990  
PCT-US91-08328-24

Query Match 26.7%; Score 4; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RRER 11  
DB 2 RRER 5

RESULT 14  
5210075-7  
Patent No. 5210075  
APPLICANT: SCHOLTZ, WOLFGANG;CHIANG, SHIU-LANG;NAGARAJAN,  
GOBI;LOBL, THOMAS J.  
TITLE OF INVENTION: INTERLEUKIN 6 ANTAGONIST PEPTIDES  
NUMBER OF SEQUENCES: 69  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,868  
FILING DATE: 16-FEB-1990  
SEQ ID NO:7  
LENGTH: 10  
5210075-7

NUMBER OF SEQUENCES: 1254  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/159,339A  
FILING DATE: 29-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,666  
FILING DATE: 07-AUG-1992  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen Lauver  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 018623-005030US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX:  
INFORMATION FOR SEQ ID NO: 135:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-159-339A-135

Query Match 26.7%; Score 4; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRE 10  
||||  
Db 4 LRRE 7

RESULT 10  
US-08-159-339A-252  
Sequence 252, Application US/08159339A  
Patent No. 6037135  
GENERAL INFORMATION:  
APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Cellis, Esteban  
TITLE OF INVENTION: HLA Binding peptides and Their  
TITLE OF INVENTION: Uses  
NUMBER OF SEQUENCES: 1254  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/159,339A  
FILING DATE: 29-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,666  
FILING DATE: 07-AUG-1992  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen Lauver  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 018623-005030US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX:  
INFORMATION FOR SEQ ID NO: 252:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-159-339A-252

Query Match 26.7%; Score 4; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRRE 10  
||||  
Db 2 LRRE 5

RESULT 11  
US-08-726-306A-1  
Sequence 1, Application US/08726306A  
Patent No. 5958684  
GENERAL INFORMATION:  
APPLICANT: van Leeuwen, Frederik Willem  
APPLICANT: Burbach, Johannes Peter Henri  
APPLICANT: Grosveld, Franklin G.  
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS  
NUMBER OF SEQUENCES: 189  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 1 Financial Center  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/726,306A  
FILING DATE: 02-Oct-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 95/20080.4  
FILING DATE: 02-Oct-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/009,832  
FILING DATE: 01-Jan-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams, Ph.D., Kathleen M.



NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-704-170-4

Query Match 26.7%; Score 4; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 RRER 11  
Db 1 RRER 4

RESULT 7  
PCT-US94-02631-4  
Sequence 4, Application PC/TUS9402631  
GENERAL INFORMATION:  
APPLICANT: Douvas, Angelina  
APPLICANT: Takehana, Yoshi  
APPLICANT: Ehresmann, Glenn  
TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
TITLE OF INVENTION: IMMUNOINFECTION CLUSTER VIRUS INFECTIONS  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Robbins, Berliner & Carson  
STREET: 201 North Figueroa Street, Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90012

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/02631  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US94-02631-4

Query Match 26.7%; Score 4; DB 5; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 RRER 11  
Db 1 RRER 4

RESULT 8  
US-09-079-447-2  
Sequence 2, Application US/09079447  
Patent No. 6083486  
GENERAL INFORMATION:  
APPLICANT: Weissleder, Ralph  
APPLICANT: Tung, Ching  
APPLICANT: Mahmood, Umar  
APPLICANT: Josephson, Lee  
APPLICANT: Bogdanov, Alexei  
TITLE OF INVENTION: INTRAMOLECULARLY-QUENCHED  
TITLE OF INVENTION: NEAR INFRARED FLUORESCENT PROBES  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA: US/09/079,447  
APPLICATION NUMBER: 08472/758001  
FILING DATE: 14-MAY-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Creason, Gary L.  
REGISTRATION NUMBER: 34,310  
REFERENCE/DOCKET NUMBER: 08472/758001  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-079-447-2

Query Match 26.7%; Score 4; DB 3; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSSK 4  
Db 1 HSSK 4

RESULT 9  
US-08-159-339A-135  
Sequence 135, Application US/08159339A  
Patent No. 6037135  
GENERAL INFORMATION:  
APPLICANT: Kudo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Esteban  
TITLE OF INVENTION: HLA Binding peptides and Their  
TITLE OF INVENTION: Uses

ADDRESSEE: David W. Clough  
STREET: Marshall, O'Toole, Gerstein, Murray & Borun  
CITY: Chicago  
STATE: Illinois  
COUNTRY: U.S.A.  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/350.260A  
FILING DATE: 05-DEC-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9110549.4  
FILING DATE: 15-MAY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9206318.9  
FILING DATE: 24-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB91/01134  
FILING DATE: 10-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB92/00883  
FILING DATE: 15-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/00605  
FILING DATE: 24-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/150,002  
FILING DATE: 31-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/307,619  
FILING DATE: 16-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 28111/32372  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
INFORMATION FOR SEQ ID NO: 415:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-350-260A-415

Query Match 26.7%; Score 4; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

Qy 7 LRRE 10  
Db 1 LRRE 4

RESULT 5  
PCT-US94-02631-14  
Sequence 14, Application PC/TUS9402631  
GENERAL INFORMATION:  
APPLICANT: Douvas, Angeline  
APPLICANT: Takehana, Yoshi  
APPLICANT: Ehresmann, Glenn  
TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Robbins, Berliner & Carson

STREET: 201 North Figueroa Street, Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/02631  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Spitals, John P.  
REGISTRATION NUMBER: 29,215  
REFERENCE/DOCKET NUMBER: 1920-331  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 977-1001  
TELEFAX: (213) 977-1003  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US94-02631-14

Query Match 26.7%; Score 4; DB 5; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0;

Qy 8 RRER 11  
Db 1 RRER 4

RESULT 6  
US-08-704-170-4  
Sequence 4, Application US/08704170  
Patent No. 5707626  
GENERAL INFORMATION:  
APPLICANT: Douvas, Angeline  
APPLICANT: Takehana, Yoshi  
APPLICANT: Ehresmann, Glenn  
TITLE OF INVENTION: IMMUNOINFECTIVE CLUSTER VIRUS INFECTIONS  
NUMBER OF SEQUENCES: 121  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Robbins, Berliner & Carson  
STREET: 201 No. 5707626th Figueroa Street, Suite 500  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/704.170  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/029,850  
FILING DATE: 11-MAR-1993  
ATTORNEY/AGENT INFORMATION:

RESULT 2  
US-08-441-871-52  
; Sequence 52, Application US/08441871  
; Patent No. 5846765  
; GENERAL INFORMATION:  
; APPLICANT: Matthews, David J.  
; APPLICANT: Wells, James A.  
; APPLICANT: Zoller, Mark J.  
; TITLE OF INVENTION: Identification of No. 5846765el Substrates  
; NUMBER OF SEQUENCES: 152  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/441,871  
; FILING DATE: 16-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/418928  
; FILING DATE: 05-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/161692  
; FILING DATE: 03-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/864452  
; FILING DATE: 06-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/09133  
; FILING DATE: 03-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/743614  
; FILING DATE: 09-AUG-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715300  
; FILING DATE: 14-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/683400  
; FILING DATE: 10-APR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/621667  
; FILING DATE: 03-DEC-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winter, Daryl B.  
; REGISTRATION NUMBER: 32,637  
; REFERENCE/DOCKET NUMBER: 645P5C2D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1249  
; TELEFAX: 415/952-9881  
; INFORMATION FOR SEQ ID NO: 52:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-441-871-52

Query Match 26.7%; Score 4; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 SSKE 5  
Db 1 SSKE 4

RESULT 3  
US-08-704-170-14  
; Sequence 14, Application US/08704170  
; Patent No. 5707626  
; GENERAL INFORMATION:  
; APPLICANT: Douvas, Angeline  
; APPLICANT: Takehana, Yoshi  
; APPLICANT: Ehresmann, Glenn  
; TITLE OF INVENTION: THERAPEUTIC STRATEGIES FOR  
; IMMUNOINFECTION: IMMUNOINFECTION CLUSTER VIRUS INFECTIONS  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Robbins, Berliner & Carson  
; STREET: 201 No. 5707626th Figueroa Street, Suite 500  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90012  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/704,170  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/029,850  
; FILING DATE: 11-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spitals, John P.  
; REGISTRATION NUMBER: 29,215  
; REFERENCE/DOCKET NUMBER: 1920-331  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 977-1001  
; TELEFAX: (213) 977-1003  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-704-170-14

Query Match 26.7%; Score 4; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RRER 11  
Db 1 RRER 4

RESULT 4  
US-08-350-260A-415  
; Sequence 415, Application US/08350260A  
; Patent No. 5962255  
; GENERAL INFORMATION:  
; APPLICANT: Winter, Gregory Paul  
; APPLICANT: Griffiths, Andrew David  
; APPLICANT: Williams, Samuel Cameron  
; APPLICANT: Waterhouse, Peter  
; APPLICANT: Nissim, Ahuva  
; APPLICANT: Johnson, Kevin Stuart  
; APPLICANT: Smith, Andrew John Hammond  
; TITLE OF INVENTION: Methods for producing members of specific  
; binding pairs  
; NUMBER OF SEQUENCES: 602  
; CORRESPONDENCE ADDRESS:

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:41:30 ; Search time 12.41 Seconds  
(without alignments)  
24.888 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_154

Perfect score: 15

Sequence: 1 HSSKEKLRRIKYC 15

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 87752

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgnl\_7/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgnl\_7/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgnl\_7/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgnl\_7/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgnl\_7/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgnl\_7/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	33.3	6	5	PCT-US93-12679-5
2	4	26.7	4	2	US-08-441-871-52
3	4	26.7	5	1	US-08-704-170-14
4	4	26.7	5	2	US-08-350-260A-415
5	4	26.7	5	5	PCT-US94-02631-14
6	4	26.7	6	1	US-08-704-170-4
7	4	26.7	6	5	PCT-US94-02631-4
8	4	26.7	7	3	US-09-079-447-2
9	4	26.7	9	3	US-08-159-339A-135
10	4	26.7	9	3	US-08-159-339A-252
11	4	26.7	10	2	US-08-726-306A-1
12	4	26.7	10	4	US-09-177-249-235
13	4	26.7	10	5	PCT-US91-08328-24
14	4	26.7	10	6	5210075-7
15	4	26.7	11	2	US-08-342-930-5
16	4	26.7	11	2	US-08-934-222-62
17	4	26.7	11	2	US-08-933-402-62
18	4	26.7	11	2	US-09-207-621-62
19	4	26.7	11	2	US-08-532-818-62
20	4	26.7	11	3	US-09-231-797-62
21	4	26.7	11	3	US-08-934-224-62
22	4	26.7	11	3	US-08-933-843-62
23	4	26.7	11	4	US-08-934-223-62
24	4	26.7	11	5	PCT-US91-08328-14
25	4	26.7	11	5	PCT-US91-08328-16
26	4	26.7	11	5	PCT-US91-08328-18
27	4	26.7	12	1	US-07-778-233B-13

28	4	26.7	12	1	US-07-963-321-13	Sequence 13, Appl
29	4	26.7	12	1	US-08-290-641-13	Sequence 13, Appl
30	4	26.7	12	1	US-08-548-540-13	Sequence 13, Appl
31	4	26.7	12	4	US-09-058-459-32	Sequence 32, Appl
32	4	26.7	12	4	US-09-127-926-32	Sequence 32, Appl
33	4	26.7	12	5	PCT-US96-09809-13	Sequence 13, Appl
34	4	26.7	13	1	US-08-452-722-1	Sequence 1, Appl
35	4	26.7	13	1	US-08-404-731A-1	Sequence 1, Appl
36	4	26.7	13	1	US-08-344-227-1	Sequence 1, Appl
37	4	26.7	13	2	US-08-503-226B-1	Sequence 1, Appl
38	4	26.7	13	3	US-08-721-458B-1	Sequence 1, Appl
39	4	26.7	14	1	US-07-909-122-4	Sequence 4, Appl
40	4	26.7	15	1	US-08-452-722-2	Sequence 2, Appl
41	4	26.7	15	1	US-08-452-722-3	Sequence 3, Appl
42	4	26.7	15	1	US-08-404-731A-2	Sequence 2, Appl
43	4	26.7	15	1	US-08-404-731A-3	Sequence 3, Appl
44	4	26.7	15	1	US-08-344-227-2	Sequence 2, Appl
45	4	26.7	15	5	PCT-US93-07545-77	Sequence 77, Appl

ALIGNMENTS

RESULT 1

PCT-US93-12679-5

; Sequence 5, Application PC/TUS9312679

; GENERAL INFORMATION:

; APPLICANT: Rath, Matthias

; TITLE OF INVENTION: SYNTHETIC OLIGOPEPTIDES ANALOGOUS TO

; TITLE OF INVENTION: PROTEIN SIGNAL SEQUENCES -METHODS OF USE

; NUMBER OF SEQUENCES: 41

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SHELDON & MAK

; STREET: 401 Florence Street

; CITY: Palo Alto

; STATE: California

; COUNTRY: USA

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/12679

; FILING DATE: 30-DEC-1993

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Cranfill, Raymond B

; REGISTRATION NUMBER: 32,845

; REFERENCE/DOCKET NUMBER: RATH-10016PC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-322-5333

; TELEFAX: 415-322-5499

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

PCT-US93-12679-5

Query Match 33.3%; Score 5; DB 5; Length 6;

Best Local Similarity 100.0%; Pred.No. 1.5e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0;

Qy 2 SSKEK 6

Db 2 SSKEK 6

RP SEQUENCE.  
RC SPECIES=H. trivolvus; TISSUE=Kidney;  
RX MEDLINE=94286417; PubMed=7912428;  
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;  
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma  
trivolvus";  
RL Peptides 15:31-36(1994).  
CC -I- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 4 4  
SQ SEQUENCE 4 AA; 582 MW; 69D40729A0000000 CRC64;

Query Match 13.3%; Score 2; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LR 8  
Db 2 LR 3

Search completed: September 15, 2001, 12:45:53  
Job time: 108 sec

ID EFIA\_MICCR STANDARD; PRT; 15 AA.  
AC P81266;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DE ELONGATION FACTOR 1-ALPHA (EF-1-ALPHA) (FRAGMENT).  
OS Microplitis croceipes.  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;  
OC Ichneumonidae; Braconidae; Microgastrinae; Microplitis.  
OX NCBI\_TaxID=72115;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=99033459; PubMed=9816671;  
RA Stuart M.K.;  
RT "An antibody diagnostic for hymenopteran parasitism is specific for a  
RT homologue of elongation factor-1 alpha.";  
RL Arch. Insect Biochem. Physiol. 39:1-8(1998).  
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF  
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN  
CC BIOSYNTHESIS (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
CC EF-TU/EF-1A SUBFAMILY.  
CC InterPro: IPR000795;  
DR PROSITE: PS00301; EFACITOR\_GTP; PARTIAL.  
KW Elongation factor; Protein biosynthesis; GTP-binding.  
FT NON\_TER 15  
SQ SEQUENCE 15 AA; 1670 MW; 1EB7DA05B09B3751 CRC64;  
  
Query Match 20.0%; Score 3; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 4 KEK 6  
|||  
Db 2 KEK 4  
  
RESULT 13  
LUXE\_VIBFI  
ID LUXE\_VIBFI STANDARD; PRT; 3 AA.  
AC P24272;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE LONG-CHAIN-FATTY-ACID--LUCIFERIN-COMPONENT LIGASE (EC 6.2.1.19)  
DE (ACYL-PROTEIN SYNTHETASE) (FRAGMENT).  
GN LUXE.  
OS Vibrio fischeri.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=668;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91072226; PubMed=2254256;  
RA Swartzman E., Kapoor S., Graham A.F., Weighen E.A.;  
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination  
RT site for the lux operon.";  
RL J. Bacteriol. 172:6797-6802(1990).  
CC -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.  
CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE  
CC FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS  
CC SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.  
CC -!- CATALYTIC ACTIVITY: ATP + AN ACID + PROTEIN = AMP + PYROPHOSPHATE  
CC + AN ACYL-PROTEIN THIOLESTER.  
CC -!- PATHWAY: SECOND STEP IN THE BIOLUMINESCENT FATTY ACID REDUCTASE  
CC COMPLEX.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M62812; ; NOT\_ANNOTATED\_CDS.  
KW Luminescence; Ligase.  
FT NON\_TER 1  
SQ SEQUENCE 3 AA; 374 MW; 6AA330303000000000 CRC64;  
  
Query Match 13.3%; Score 2; DB 1; Length 3;  
Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 12 IK 13  
||  
Db 1 IK 2  
  
RESULT 14  
FAR3\_HIRME  
ID FAR3\_HIRME STANDARD; PRT; 4 AA.  
AC P42562;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE FMRFAMIDE-LIKE NEUROPEPTIDE YLRP-AMIDE.  
OS Hirudo medicinalis (Medicinal leech).  
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;  
OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.  
OX NCBI\_TaxID=6421;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92195954; PubMed=1686933;  
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;  
RT "Identification of RFamide neuropeptides in the medicinal leech.";  
RL Peptides 12:897-908(1991).  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
CC Neuropeptide; Amidation.  
KW MOD\_RES 4  
FT SEQUENCE 4 AA; 598 MW; 69D4073B300000000 CRC64;  
SQ SEQUENCE 4 AA; 598 MW; 69D4073B300000000 CRC64;  
  
Query Match 13.3%; Score 2; DB 1; Length 4;  
Best Local Similarity 100.0%; Pred. No. 9.3e+04;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 7 LR 8  
||  
Db 2 LR 3  
  
RESULT 15  
FLRF\_HIRME  
ID FLRF\_HIRME STANDARD; PRT; 4 AA.  
AC P42561;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE FLRFAMIDE.  
OS Hirudo medicinalis (Medicinal leech), and Helisoma trivolvis (Snail).  
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;  
OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.  
OX NCBI\_TaxID=6421, 27815;  
RN [1]  
RP SEQUENCE.  
RX SPECIES-H medicinalis;  
RX MEDLINE=92195954; PubMed=1686933;  
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;  
RT "Identification of RFamide neuropeptides in the medicinal leech.";  
RL Peptides 12:897-908(1991).  
RN [2]

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 KYC 15

Db 9 KYC 11

RESULT 9

UR2\_SCYCA ID UR2\_SCYCA STANDARD; PRT; 12 AA.

AC P35490;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE UROTENSIN II (U-II) (UII).

OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;

OC Elasmobranchii; Galeomorphi; Galeoidea; Carcharhiniformes;

OC Scyllorhinidae; Scyliorhinus.

OX NCBI\_TaxID=7830;

RN [1]

RP SEQUENCE.

RC TISSUE=Spinal cord;

RX MEDLINE=92319231; PubMed=1620290;

RA Conlon J.M., O'Harte F., Smith B.D., Balment R.J., Hazon N.;

RT "Purification and characterization of urotensin II and parvalbumin

from an elasmobranch fish, Scyliorhinus canicula (common dogfish).";

RL Neuroendocrinology 55:230-235(1992).

CC -1- FUNCTION: HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A

MUSCLE STIMULATION.

CC -1- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.

DR InterPro; IPR001483; -

DR Pfam; PF02083; Urotensin II; 1.

DR PROSITE; PS00984; UROTENSIN\_II; 1.

KW Hormone.

FT DISULFID 6 11

SQ SEQUENCE 12 AA; 1526 MW; 804729F9D579CEBA CRC64;

Query Match

Best Local Similarity 20.0%; Score 3; DB 1; Length 12;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 KYC 15

Db 9 KYC 11

RESULT 10

IDHP\_RAT

AC P56574; STANDARD; PRT; 13 AA.

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE ISOCITRATE DEHYDROGENASE [NADP], MITOCHONDRIAL (EC 1.1.1.42)

DE (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC ICDH) (IDP) (ICD-

M) (FRAGMENT).

GN IDH2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE

RC STRAIN=WISTAR; TISSUE=Heart;

RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,

RA Jungblut P.R.;

RL Submitted (SEP-1998) to the SWISS-PROT data bank.

CC -1- FUNCTION: PLAYS A ROLE IN INTERMEDIARY METABOLISM AND ENERGY

PRODUCTION. IT MAY TIGHTLY ASSOCIATE OR INTERACT WITH THE PYRUVATE

DEHYDROGENASE COMPLEX (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: ISOCITRATE + NADP(+) -> 2-OXOGLUTARATE +

CO(2) + NADPH.

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.

CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN

(SPOT P8) IS: 9.0. ITS MW IS: 42 KDA.

CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE AND ISOPROPYLMALATE

DEHYDROGENASES FAMILY.

DR InterPro; IPR001804; -

DR PROSITE; PS00470; IDH, IMDH; PARTIAL.

KW Oxidoreductase; NADP; Glyoxylate bypass; Tricarboxylic acid cycle;

KW Mitochondrion

FT NON\_TER 13 13

SQ SEQUENCE 13 AA; 1526 MW; 5FB81031723E02C3 CRC64;

Query Match

Best Local Similarity 20.0%; Score 3; DB 1; Length 13;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RIK 13

Db 4 RIK 6

RESULT 11

CAT2\_FASHE

ID CAT2\_FASHE STANDARD; PRT; 14 AA.

AC P80342;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE CATHEPSIN L2 (EC 3.4.22.15) (FRAGMENT).

OS Fasciola hepatica (Liver fluke).

OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;

OC Trematoda; Digenea; Echinostomida; Echinostomata; Fascioloidae;

OC Fasciolidae; Fasciola.

OX NCBI\_TaxID=6192;

RN [1]

RP SEQUENCE.

RX MEDLINE=94307282; PubMed=8033913;

RA Dowd A.J., Smith A.M., McConicle S., Dalton J.P.;

RT "Purification and characterisation of a second cathepsin L proteinase

secreted by the parasitic trematode Fasciola hepatica.";

RL Eur. J. Biochem. 223:91-98(1994).

CC -1- FUNCTION: THIOL PROTEASE THAT ASSISTS THE PARASITE IN BURROWING

THROUGH THE GUT WALL AND LIVER OF ITS MAMMALIAN HOST.

CC -1- SUBUNIT: DIMER OF AN HEAVY AND A LIGHT CHAIN LINKED BY DISULFIDE

BONDS.

CC -1- SUBCELLULAR LOCATION: LYSOSOMAL.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE

PAPAIN FAMILY OF THIOL PROTEASES.

DR MEROPS: C01.033; -

DR InterPro; IPR000169; -

DR PROSITE; PS00139; THIOL\_PROTEASE\_CYS; PARTIAL.

DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; PARTIAL.

DR PROSITE; PS00640; THIOL\_PROTEASE\_ASN; PARTIAL.

KW Hydrolase; Thiol protease; Lysosome.

FT NON\_TER 14 14

SQ SEQUENCE 14 AA; 1605 MW; 9CAEAB74E9D4110A CRC64;

Query Match

Best Local Similarity 20.0%; Score 3; DB 1; Length 14;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RRE 10

Db 8 RRE 10

RESULT 12

EFIA\_MICCR

RT "Isolation and amino acid sequence of two urotensin II peptides from  
 RL Catostomus commersoni urophyses.";   
 CC Peptides 4:367-373(1983)   
 CC -1- FUNCTION: UROSENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY   
 CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A   
 CC CORTICOTROPIN-RELEASING FACTOR.   
 CC -1- SIMILARITY: BELONGS TO THE UROSENSIN 2 FAMILY.   
 CC PIR: JS0424; JS0424.   
 DR InterPro: IPR001483; -   
 DR Pfam: PF02083; Urotensin\_II; 1.   
 DR PROSITE: PS00984; UROSENSIN\_II; 1.   
 KW Hormone.   
 FT DISULFID 6 11   
 SQ SEQUENCE 12 AA; 1437 MW; 73961BDBB879CEBB CRC64;

Query Match 20.0%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 KYC 15  
 Db 9 KYC 11

RESULT 6  
 UR2B\_CYPCA STANDARD; PRT; 12 AA.  
 AC P04561;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE UROSENSIN II-BETA.  
 OS Cyprinus carpio (Common carp).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 CC Cypriniformes; Cyprinidae; Cyprininae; Cyprinus.  
 OX NCBI\_TaxID=7962;  
 RN [1]  
 RP SEQUENCE.  
 RA Muneata E., Ohtaki T., Ichikawa T., McMaster D., Lederis K.;  
 RL (In) Rich D.H., Gross E. (eds.);  
 RL Proceedings of the 7th American peptide symposium, pp.69-72.  
 RL Pierce Chemical Co., Rockford IL (1981).  
 CC -1- FUNCTION: UROSENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY   
 CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A   
 CC CORTICOTROPIN-RELEASING FACTOR.   
 CC -1- SIMILARITY: BELONGS TO THE UROSENSIN 2 FAMILY.   
 CC InterPro: IPR001483; -   
 DR Pfam: PF02083; Urotensin\_II; 1.   
 DR PROSITE: PS00984; UROSENSIN\_II; 1.   
 KW Hormone.   
 FT DISULFID 6 11   
 FT VARIANT 2 2 G -> S.   
 SQ SEQUENCE 12 AA; 1407 MW; 73960A9FB879CEBB CRC64;

Query Match 20.0%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 KYC 15  
 Db 9 KYC 11

RESULT 7  
 UR2\_GILMI STANDARD; PRT; 12 AA.  
 AC P01147;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE UROSENSIN II (U-II) (UII).  
 OS Gillichthys mirabilis (Long-jawed mudsucker).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Gobioidel;  
 CC Gobiidae; Gillichthys.  
 OX NCBI\_TaxID=8222;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=81054904; PubMed=6107911;  
 RA Pearson D., Shively J.E., Clark B.R., Geschwind I.I., Barkley M.,  
 RA Nishioka R., Bern H.A.;  
 RT "Urotensin II: a somatostatin-like peptide in the caudal  
 RT neurosecretory system of fishes.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 77:5021-5024(1980).  
 CC -1- FUNCTION: UROSENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY   
 CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A   
 CC CORTICOTROPIN-RELEASING FACTOR.   
 CC -1- SIMILARITY: BELONGS TO THE UROSENSIN 2 FAMILY.   
 CC PIR: A01409; UOCM2.  
 DR InterPro: IPR001483; -   
 DR Pfam: PF02083; Urotensin\_II; 1.   
 DR PROSITE: PS00984; UROSENSIN\_II; 1.   
 KW Hormone.   
 FT DISULFID 6 11   
 SQ SEQUENCE 12 AA; 1364 MW; 968BF8982679CEBA CRC64;

Query Match 20.0%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 KYC 15  
 Db 9 KYC 11

RESULT 8  
 UR2\_POLSP STANDARD; PRT; 12 AA.  
 ID UR2\_POLSP  
 AC P81022;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE UROSENSIN II (U-II) (UII).  
 OS Polyodon spathula (North American paddlefish).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Chondrostei; Acipenseriformes; Polyodontidae;  
 CC Polyodon.  
 OX NCBI\_TaxID=7913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spinal cord;  
 RX MEDLINE=96051494; PubMed=8536944;  
 RA Waugh D., Youson J., Mims S.D., Sower S., Conlon J.M.;  
 RT "Urotensin II from the river lamprey (Lampetra fluviatilis), the sea  
 RT lamprey (Petromyzon marinus), and the paddlefish (Polyodon  
 RT spathula).";   
 RL Gen. Comp. Endocrinol. 99:323-332(1995).  
 CC -1- FUNCTION: HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A   
 CC CORTICOTROPIN-RELEASING FACTOR. PROBABLY INVOLVED IN SMOOTH   
 CC MUSCLE STIMULATION.   
 CC -1- SIMILARITY: BELONGS TO THE UROSENSIN 2 FAMILY.   
 CC InterPro: IPR001483; -   
 DR Pfam: PF02083; Urotensin\_II; 1.   
 DR PROSITE: PS00984; UROSENSIN\_II; 1.   
 KW Hormone.   
 FT DISULFID 6 11 BY SIMILARITY.   
 SQ SEQUENCE 12 AA; 1410 MW; 7551E9DBB879CEBB CRC64;

Query Match 20.0%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;



GN RPS16.  
 OS Ginkgo biloba (Ginkgo).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Ginkgoophyta; Ginkgoales; Ginkgoaceae; Ginkgo.  
 OX NCBI\_TaxID=3311;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95094313; PubMed=8001171;  
 RA Richard M., Tremblay C., Bellemare G.;  
 RT "Chloroplastic genomes of Ginkgo biloba and Chlamydomonas moewusii  
 contain a chlB gene encoding one subunit of a light-independent  
 protochlorophyllide reductase.";  
 RL Curr. Genet. 26:159-165(1994).  
 CC -!- SIMILARITY: BELONGS TO THE S16P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U01531; AAA6977.1; -;  
 DR InterPro; IPR00307; -;  
 DR PROSITE; PS00732; RIBOSOMAL\_S16; 1.  
 KW Ribosomal protein; Chloroplast.  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1488 MW; 6700EDAF9D033734 CRC64;  
 Query Match 20.0%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 KLR 8  
 DB 3 KLR 5  
 RESULT 3  
 TM2A\_METMA STANDARD; PRT; 12 AA.  
 ID TM2A\_METMA  
 AC P80652;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE ALTERNATIVE TETRAHYDROMETHANOPTERIN S-METHYLTRANSFERASE 28 KDA SUBUNIT  
 DE (EC 2.1.1.86) (N5-METHYLTETRAHYDROMETHANOPTERIN--COENZYME M  
 DE METHYLTRANSFERASE 28 KDA SUBUNIT) (FRAGMENT).  
 OS Methanosarcina mazel.  
 OS Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;  
 OC Methanosarcina.  
 OX NCBI\_TaxID=2209;  
 [1]  
 RP SEQUENCE.  
 RC STRAIN=DSM 3647 / GOE1;  
 RX MEDLINE=96370840; PubMed=8774736;  
 RA Lienard T., Becher B., Marschall M., Bowen S., Gottschalk G.;  
 RT "Sodium ion translocation by N5-methyltetrahydromethanopterin:  
 RT coenzyme M methyltransferase from Methanosarcina mazel Gol  
 RT reconstituted in ether lipid liposomes.";  
 RL Eur. J. Biochem. 239:857-864(1996).  
 CC -!- FUNCTION: THIS ENZYME COMPLEX CATALYZES AN INTERMEDIATE STEP IN  
 CC METHANOGENESIS, THE FORMATION OF METHYL-COENZYME M AND  
 CC TETRAHYDROMETHANOPTERIN.  
 CC TETRAHYDROMETHANOPTERIN.  
 CC -!- CATALYTIC ACTIVITY: 5-METHYL-5,6,7,8-TETRAHYDROMETHANOPTERIN +  
 CC 2-MERCAPTOETHANESULFONATE = 5,6,7,8-TETRAHYDROMETHANOPTERIN +  
 CC 2-(METHYLTHIO)ETHANESULFONATE.  
 CC -!- SUBUNIT: COMPOSED OF SIX DIFFERENT SUBUNITS.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

KW Transferase; Methyltransferase; Transmembrane; Methanogenesis.  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1321 MW; 6DEAA5766232D76B CRC64;  
 Query Match 20.0%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 EKL 7  
 DB 2 EKL 4  
 RESULT 4  
 UR2A\_CATCO STANDARD; PRT; 12 AA.  
 ID UR2A\_CATCO  
 AC P04558;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE UROTENSIN IIA (U-IIA) (UIIA).  
 DE CATOSTOMUS COMMERSONI (White sucker).  
 OS Catostomus commersoni (White sucker).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Catostomidae; Catostomus.  
 OX NCBI\_TaxID=7971;  
 [1]  
 RP SEQUENCE.  
 RX MEDLINE=84041959; PubMed=6138758;  
 RA McMaster D., Lederis K.;  
 RT "Isolation and amino acid sequence of two urotensin II peptides from  
 RT Catostomus commersoni urophyses.";  
 RL Peptides 4:367-373(1983).  
 CC -!- FUNCTION: UROTENSIN IS FOUND IN THE TELEOST CAUDAL NEUROSECRETORY  
 CC SYSTEM. IT HAS A SUGGESTED ROLE IN OSMOREGULATION AND AS A  
 CC CORTICOTROPIN-RELEASING FACTOR.  
 CC -!- SIMILARITY: BELONGS TO THE UROTENSIN 2 FAMILY.  
 DR PIR; JS0423; JS0423.  
 DR InterPro; IPR001483; -;  
 DR Pfam; PF02083; Urotensin\_II; 1.  
 DR PROSITE; PS00984; UROTENSIN\_II; 1.  
 KW Hormone.  
 FT DISULFID 6 11  
 SQ SEQUENCE 12 AA; 1336 MW; 969C76DBB879CEBA CRC64;  
 Query Match 20.0%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 13 KYC 15  
 DB 9 KYC 11  
 RESULT 5  
 UR2B\_CATCO STANDARD; PRT; 12 AA.  
 ID UR2B\_CATCO  
 AC P04559;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE UROTENSIN IIB (U-IIB) (UIIB).  
 OS Catostomus commersoni (White sucker).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Catostomidae; Catostomus.  
 OX NCBI\_TaxID=7971;  
 [1]  
 RP SEQUENCE.  
 RX MEDLINE=84041959; PubMed=6138758;  
 RA McMaster D., Lederis K.;

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:44:05 ; Search time 10.36 Seconds  
(without alignments)  
49.596 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_154  
Perfect score: 15  
Sequence: 1 HSSKEKLRRRIKYC 15

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 699

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3	20.0	5	1 BIOB_CITFR	P12997 citrobacter
2	3	20.0	12	1 RR16_GINBI	P36207 ginkgo bilo
3	3	20.0	12	1 TM2A_METMA	P80652 methanosarc
4	3	20.0	12	1 UR2A_CATCO	P04558 catostomus
5	3	20.0	12	1 UR2B_CATCO	P04559 catostomus
6	3	20.0	12	1 UR2B_CYPCA	P04561 cyprinus ca
7	3	20.0	12	1 UR2_GILMI	P01147 gyllichthys
8	3	20.0	12	1 UR2_POLSP	P81022 polyodon sp
9	3	20.0	12	1 UR2_SCYCA	P35490 scyllorhinu
10	3	20.0	13	1 IDHP_RAT	P56574 rattus norv
11	3	20.0	14	1 CAT2_FASHE	P80342 fasciola he
12	3	20.0	15	1 EF1A_MICCR	P81266 microplitis
13	2	13.3	3	1 LUXE_VIBFI	P24272 vibrio fisc
14	2	13.3	4	1 FAR3_HIRME	P42562 hirudo medi
15	2	13.3	4	1 FLRF_HIRME	P42561 hirudo medi
16	2	13.3	6	1 ACPH_RABIT	P25154 oryctolagus
17	2	13.3	6	1 LOK1_LOCHI	P41491 locusta mig
18	2	13.3	7	1 CARP_MYTED	P10420 mytilus edu
19	2	13.3	7	1 CHOX_ALCSP	P16101 alcaligenes
20	2	13.3	7	1 FAR1_HELTI	P41871 helisoma tr
21	2	13.3	7	1 FAR1_PROCL	P38499 procambarus
22	2	13.3	7	1 FAR2_ASCSU	P31890 ascaris suu
23	2	13.3	7	1 FAR2_PROCL	P38498 procambarus
24	2	13.3	7	1 FAR5_HIRME	P42564 hirudo medi
25	2	13.3	7	1 HY7_PIG	P01153 sus scrofa
26	2	13.3	7	1 MSCL_SALTY	P39446 salmonella
27	2	13.3	7	1 MYOM_APLCA	P15513 aplysia cal
28	2	13.3	7	1 UH11_RAT	P56576 rattus norv
29	2	13.3	7	1 WWAL_ACHFU	P35919 achatina fu
30	2	13.3	7	1 WWAS_ACHFU	P35921 achatina fu
31	2	13.3	8	1 AC1_THUAL	P18691 thunnus alb
32	2	13.3	8	1 FAR1_PANRE	P41872 panagrellus
33	2	13.3	8	1 FAR3_HOMAM	P41486 homarus ame

34	2	13.3	8	1 FAR4_HOMAM	P41487 homarus ame
35	2	13.3	8	1 GLUR_HUMAN	P02729 homo sapien
36	2	13.3	8	1 LCK2_LEUNA	P21141 leucophaea
37	2	13.3	8	1 LCK4_LEUNA	P21143 leucophaea
38	2	13.3	8	1 LCK5_LEUNA	P19987 leucophaea
39	2	13.3	8	1 LCK6_LEUNA	P19988 leucophaea
40	2	13.3	8	1 LCK7_LEUNA	P19989 leucophaea
41	2	13.3	8	1 NS3_MYCTU	P81152 mycobacteri
42	2	13.3	8	1 RS10_SALTY	O68928 salmonella
43	2	13.3	8	1 UF06_MOUSE	P38644 mus musculu
44	2	13.3	8	1 UH09_RAT	P56575 rattus norv
45	2	13.3	9	1 BUK_CLOPA	P81337 clostridium

ALIGNMENTS

RESULT 1

ID	BIOB_CITFR	STANDARD;	PRT;	5 AA.
AC	P12997;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	BIOBIN SYNTHASE (EC 2.8.1.6) (BIOTIN SYNTHETASE) (FRAGMENT).			
GN	BIOB.			
OS	Citrobacter freundii.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Citrobacter.			
OX	NCBI_TaxID=546;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89006280; PubMed=2971595;			
RA	Shuan D., Campbell A.;			
RT	"Transcriptional regulation and gene arrangement of Escherichia coli,			
RT	Citrobacter freundii and Salmonella typhimurium biotin operons.";			
RL	Gene 67:203-211(1988).			
CC	-1- CATALYTIC ACTIVITY: DETHIOBIOTIN + (S) - BIOTIN.			
CC	-1- PATHWAY: LAST STEP IN BIOTIN BIOSYNTHESIS PATHWAY.			
CC	-1- SIMILARITY: BELONGS TO THE BIOTIN AND LIPOIC ACID SYNTHETASES FAMILY.			

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; M21922; -; NOT\_ANNOTATED\_CDS.  
KW Biotin biosynthesis; Iron-sulfur; Transferase.  
FT NON\_TER 5  
SQ SEQUENCE 5 AA; 532 MW; 75A5B1EDD6F00000 CRC64;

Query Match 20.0%; Score 3; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred.No. 9.3e+04;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 HSS 3  
Db 3 HSS 5

RESULT 2

ID	RR16_GINBI	STANDARD;	PRT;	12 AA.
AC	P36207;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	01-FEB-1996 (Rel. 33, Last annotation update)			
DE	CHLOROPLAST 30S RIBOSOMAL PROTEIN S16 (FRAGMENT).			

## RESULT 14

S42765  
urotensin II - telostean fish  
C:Species: telostean fish  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 05-Dec-1997  
C:Accession: S42765  
R:Bhaskaran, R.; Arunkumar, A.I.; Yu, C.  
Biochim. Biophys. Acta 1199, 115-122, 1994  
A:Title: NMR and dynamical simulated annealing studies on the solution conformation of u  
A:Reference number: S42765; MUID:94169160  
A:Accession: S42765  
A:Molecule type: protein  
A:Residues: 1-12 <BHA>  
C:Superfamily: urotensin II  
C:Keywords: neuropeptide; osmoregulation  
F:6-11/Product: urotensin II #status experimental <MAT>

Query Match 20.0%; Score 3; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 3.8e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 KYC 15  
|||  
Db 9 KYC 11

## RESULT 15

JS0423  
urotensin II-A peptide - white sucker  
C:Species: Catostomus commersoni (white sucker)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-May-1997  
C:Accession: JS0423  
R:McMaster, D.; Lederis, K.  
Peptides 4, 367-373, 1983  
A:Title: Isolation and amino acid sequence of two urotensin II peptides from Catostomus  
A:Reference number: JS0423; MUID:84041959  
A:Accession: JS0423  
A:Molecule type: protein  
A:Residues: 1-12 <MCW>  
C:Comment: This peptide has smooth muscle-stimulating activity.  
C:Superfamily: urotensin II  
F:6-11/Disulfide bonds: #status experimental

Query Match 20.0%; Score 3; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 3.8e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 KYC 15  
|||  
Db 9 KYC 11

Search completed: September 15, 2001, 12:44:48  
Job time: 264 sec

Query Match 20.0%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KEK 6  
Db 3 KEK 5

RESULT 9  
PT0081  
C:Species: Arabidopsis thaliana (fragment)  
C:Date: 07-Feb-1996 #sequence\_revision 19-Apr-1996 #text\_change 24-Nov-1999  
C:Accession: PT0081  
R:Tsugita, A.; Kamo, M.; Kawakami, M.; Ohki, Y.  
submitted to JIPID, December 1995  
A:Description: Two dimensional electrophoresis of plant proteins and standardization of  
A:Reference number: PN0173  
A:Accession: PT0081  
A:Molecule type: protein  
A:Residues: 1-11 <TSU>  
A:Experimental source: Leaf  
C:Keywords: acetylated amino end  
F:1/Modified site: acetylated amino end (Ser) #status experimental

Query Match 20.0%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SKE 5  
Db 1 SKE 3

RESULT 10  
SI9775  
wound-induced protein - tomato (fragment)  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Sep-1997  
C:Accession: SI9775  
R:Parsons, B.L.  
submitted to the EMBL Data Library, May 1991  
A:Reference number: SI9773  
A:Accession: SI9775  
A:Molecule type: mRNA  
A:Residues: 1-11 <PAR>  
A:Cross-references: EMBL:X59884; NID:g19323; PID:g19324

Query Match 20.0%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSK 4  
Db 3 SSK 5

RESULT 11  
I54193  
Rhesus blood group CcEe protein - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: I54193  
R:Cherif-Zahar, B.; Le Van Kim, C.; Rouillac, C.; Raynal, V.; Carttron, J.P.; Collin, Y.  
Genomics 19, 68-74, 1994  
A:Title: Organization of the gene (RHCE) encoding the human blood group RHCCe antigens  
A:Reference number: I54193; MUID:94245182  
A:Accession: I54193  
A:Status: preliminary; translated from GB/EMBL/DBJ

Query Match 20.0%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KEK 6  
Db 3 KEK 5

RESULT 9  
PT0081  
C:Species: Arabidopsis thaliana (fragment)  
C:Date: 07-Feb-1996 #sequence\_revision 19-Apr-1996 #text\_change 24-Nov-1999  
C:Accession: PT0081  
R:Tsugita, A.; Kamo, M.; Kawakami, M.; Ohki, Y.  
submitted to JIPID, December 1995  
A:Description: Two dimensional electrophoresis of plant proteins and standardization of  
A:Reference number: PN0173  
A:Accession: PT0081  
A:Molecule type: protein  
A:Residues: 1-11 <TSU>  
A:Experimental source: Leaf  
C:Keywords: acetylated amino end  
F:1/Modified site: acetylated amino end (Ser) #status experimental

Query Match 20.0%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SKE 5  
Db 1 SKE 3

RESULT 10  
SI9775  
wound-induced protein - tomato (fragment)  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Sep-1997  
C:Accession: SI9775  
R:Parsons, B.L.  
submitted to the EMBL Data Library, May 1991  
A:Reference number: SI9773  
A:Accession: SI9775  
A:Molecule type: mRNA  
A:Residues: 1-11 <PAR>  
A:Cross-references: EMBL:X59884; NID:g19323; PID:g19324

Query Match 20.0%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSK 4  
Db 3 SSK 5

RESULT 11  
I54193  
Rhesus blood group CcEe protein - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: I54193  
R:Cherif-Zahar, B.; Le Van Kim, C.; Rouillac, C.; Raynal, V.; Carttron, J.P.; Collin, Y.  
Genomics 19, 68-74, 1994  
A:Title: Organization of the gene (RHCE) encoding the human blood group RHCCe antigens  
A:Reference number: I54193; MUID:94245182  
A:Accession: I54193  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-11 <RES>  
A:Cross-references: GB:S70456; NID:g546795; PIDN:AAD14061.1; PID:g4261761  
C:Genetics:  
A:Gene: GDB:RHCE  
A:Cross-references: GDB:229957; OMIM:111700  
A:Map position: 1p36.2-1p34

Query Match 20.0%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSK 4  
Db 2 SSK 4

RESULT 12  
PH0939  
T-cell receptor beta chain V-D-J region (clone 10) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
C:Accession: PH0939  
R:Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
J. Exp. Med. 174, 1467-1476, 1991  
A:Title: Analysis of T cell receptor beta chains in Lewis rats with experimental alle  
A:Reference number: PH0891; MUID:92078857  
A:Accession: PH0939  
A:Molecule type: mRNA  
A:Residues: 1-11 <GOL>  
A:Experimental source: complete Freund's adjuvant-immunized lymph node  
C:Keywords: T-cell receptor

Query Match 20.0%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSK 4  
Db 3 SSK 5

RESULT 13  
U06M2  
urotensin II - long-jawed mudsucker  
C:Species: Gillichthys mirabilis (long-jawed mudsucker)  
C:Date: 23-Oct-1981 #sequence\_revision 23-Oct-1981 #text\_change 15-Oct-1996  
C:Accession: A01409  
R:Pearson, D.; Shively, J.E.; Clark, B.R.; Geschwind, I.I.; Barkley, M.; Nishioaka, R.  
Proc. Natl. Acad. Sci. U.S.A. 77, 5021-5024, 1980  
A:Title: Urotensin II: a somatostatin-like peptide in the caudal neurosecretory syste  
A:Reference number: A01409; MUID:81054904  
A:Accession: A01409  
A:Molecule type: protein  
A:Residues: 1-12 <PEA>  
A:Note: the proposed sequence was confirmed by synthesis of a peptide with the same s  
C:Comment: Urotensin II is found in the teleost caudal neurosecretory system and is 1  
C:Superfamily: urotensin II  
C:Keywords: neuropeptide; osmoregulation  
F:6-11/Disulfide bonds: #status experimental

Query Match 20.0%; Score 3; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 3.8e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 KYC 15  
Db 9 KYC 11

```

RESULT 3
140698
biotin B - Citrobacter freundii (fragment)
C:Species: Citrobacter freundii
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 12-Aug-1996
C:Accession: I40698
R:Shiuan, D.; Campbell, A.
Gene 67, 203-211, 1988
A:Title: Transcriptional regulation and gene arrangement of Escherichia coli, Citrobacter
A:Reference number: I40697; MUID:89006280
A:Accession: I40698
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5 <RES>
A:Cross-references: GB:M21922; NID:g144434

Query Match      20.0%; Score 3; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSS 3
   |||
Db 3 HSS 5

RESULT 4
A11490
Pyruvate kinase (EC 2.7.1.40) - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 03-Mar-1995
C:Accession: A11490
R:Hjeltnquist, G.; Andersson, J.; Edlund, B.; Engstrom, L.
Biochem. Biophys. Res. Commun. 61, 559-563, 1974
A:Title: Amino acid sequence of a (32-P)phosphopeptide from pig liver pyruvate kinase ph
A:Reference number: A11490; MUID:75127438
A:Accession: A11490
A:Molecule type: protein
A:Residues: 1-6 <HJE>
A:Experimental source: liver
A:Keywords: glycolysis; phosphotransferase

Query Match      20.0%; Score 3; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRR 9
   |||
Db 1 LRR 3

RESULT 5
FS0254
18K protein 5507 - rice (strain Nihonbare) (fragment)
C:Species: Oryza sativa (rice)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Mar-1995
C:Accession: FS0254
R:Tsuigita, A.
submitted to JIPID, April 1993
A:Reference number: FS0206
A:Accession: FS0254
A:Molecule type: protein
A:Residues: 1-7 <TSU>
A:Experimental source: leaf, chloroplast, strain Nihonbare
A:Note: molecular weight 18K, pI 4.4

Query Match      20.0%; Score 3; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSS 3
   |||
Db 3 HSS 5

RESULT 6
PT0323
Ig heavy chain CDR3 region (clone J2-106B) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0323
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an
A:Reference number: PT0222; MUID:91108337
A:Accession: PT0323
A:Molecule type: DNA
A:Residues: 1-8 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match      20.0%; Score 3; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RRE 10
   |||
Db 6 RRE 8

RESULT 7
PT0231
Ig heavy chain CDR3 region (clone 1-118B) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0231
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an
A:Reference number: PT0222; MUID:91108337
A:Accession: PT0231
A:Molecule type: DNA
A:Residues: 1-9 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match      20.0%; Score 3; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSS 3
   |||
Db 3 HSS 5

RESULT 8
H61308
hemocyanin subunit IV - Atlantic horseshoe crab (fragment)
C:Species: Limulus polyphemus (Atlantic horseshoe crab)
C:Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
C:Accession: H61308
R:Jolles, J.; Jolles, P.; Lamy, J.; Lamy, J.
FEBS Lett. 106, 289-291, 1979
A:Title: Structural characterization of seven different subunits in Androctonus austr
A:Reference number: A61308; MUID:80047238
A:Accession: H61308
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <JOL>
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:40:24 ; Search time 12.42 Seconds  
(without alignments)  
91.998 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_154  
Perfect score: 15  
Sequence: 1 HSSKEKLRRRIKYC 15

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2511

Minimum DB seq length: 0  
Maximum DB seq length: 15

Post-processing: Listing first 45 summaries

Database : PIR\_68:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4	26.7	6	B33932	Ig mu chain D region (D23) - mouse
2	4	26.7	14	PC4382	dehydrin 4.5K poly
3	3	20.0	5	I40698	biotin B - Citroba
4	3	20.0	6	A11490	pyruvate kinase (E
5	3	20.0	7	PS0254	18K protein 5507 -
6	3	20.0	8	PT0323	Ig heavy chain CRD
7	3	20.0	9	PT0231	Ig heavy chain CDR
8	3	20.0	10	H61308	hemocyanin subunit
9	3	20.0	11	PT0081	protein QA300023 -
10	3	20.0	11	S19775	wound-induced prot
11	3	20.0	11	I54193	Rhesus blood group
12	3	20.0	11	PH0939	T-cell receptor be
13	3	20.0	12	U06M2	urotensin II - lon
14	3	20.0	12	S42765	urotensin II - tel
15	3	20.0	12	SJ0423	urotensin II-A pep
16	3	20.0	12	S26558	T-cell receptor be
17	3	20.0	12	S01222	translation elonga
18	3	20.0	12	A58501	24K kidney and bia
19	3	20.0	12	T44420	hypothetical prote
20	3	20.0	12	SJ0424	urotensin II-B pep
21	3	20.0	13	PQ0445	urotensin II - lau
22	3	20.0	13	S78519	ribosomal protein
23	3	20.0	13	JQ2309	hypothetical 1.6K
24	3	20.0	13	JQ2319	hypothetical 1.6K
25	3	20.0	13	S57571	T cell receptor al
26	3	20.0	13	PH0138	T-cell receptor be
27	3	20.0	13	PH1595	Ig H chain V-D-J r
28	3	20.0	14	C33098	223K exoantigen -
29	3	20.0	14	S45655	cathepsin L (EC 3.

30	3	20.0	14	2	PC1215	homeotic protein E
31	3	20.0	14	2	A61308	hemocyanin chain 2
32	3	20.0	14	2	B61308	hemocyanin chain 4
33	3	20.0	14	4	S00843	hypothetical prote
34	3	20.0	15	2	A47146	topoisomerase I -
35	3	20.0	15	2	S61284	phosphoprotein, 80
36	3	20.0	15	2	PA0053	protein QF200006 -
37	3	20.0	15	2	S57584	T cell receptor V-
38	3	20.0	15	2	F44823	synaptosomal-assoc
39	3	20.0	15	2	PH1590	Ig H chain V-D-J r
40	3	20.0	15	2	PT0094	succinate dehydrog
41	3	20.0	15	2	FX0031	mixed lymphocyte r
42	3	20.0	15	2	A56049	urinary tract ston
43	3	20.0	15	4	I38032	hypothetical MNL/T
44	3	20.0	15	4	I38031	hypothetical MNL/T
45	2	13.3	3	2	PT0622	T-cell receptor be

ALIGNMENTS

RESULT 1

-B33932  
Ig mu chain D region (D23) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 16-Aug-1996  
C:Accession: B33932  
R:Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.  
Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989  
A:Title: Two murine natural polyreactive autoantibodies are encoded by nonmutated ger  
A:Reference number: A33932; MUID:89282823  
A:Accession: B33932  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-6 <BAC>  
A:Cross-references: GB:M27107  
C:Keywords: immunoglobulin

Query Match 26.7%; Score 4; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.2e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 EKLK 8  
Db 1 EKLK 4

RESULT 2

PC4382  
dehydrin 4.5K polypeptide - Soybean (fragment)  
N:Alternate names: acid soluble 26K protein  
C:Species: Glycine max cv. Nattosoryu (soybean)  
C:Date: 06-Nov-1997 #sequence\_revision 06-Nov-1997 #text\_change 07-May-1999  
C:Accession: PC4382  
R:Momma, M.; Harauchi, K.; Saito, M.; Chikuni, K.; Harada, K.  
Biosci. Biotechnol. Biochem. 61, 1286-1291, 1997  
A:Title: Purification and characterization of the acid soluble 26-kDa polypeptide fro  
A:Reference number: PC4380; MUID:97446521  
A:Accession: PC4382  
A:Molecule type: protein  
A:Residues: 1-14 <MOM>  
A:Experimental source: seed  
C:Comment: This protein is rich in hydrophilic amino acid, and highly heat stable.

Query Match 26.7%; Score 4; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 KEKL 7  
Db 5 KEKL 8

---

RT transcriptional regulator, DNR, in addition to ANR.";  
 RL FEBS Lett. 371:73-76(1995).  
 DR EMBL; D50019; BAA08746.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 11 AA; 1543 MW; DF363CAE141B5736 CRC64;

Query Match 20.0%; Score 3; DB 2; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 8.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRR 9  
 |||  
 Db 9 LRR 11

RESULT 14  
 Q9K332 PRELIMINARY; PRT; 11 AA.  
 AC Q9K332;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE GEH (FRAGMENT).  
 GN GEH.  
 OS Staphylococcus aureus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Staphylococcus.  
 OX NCBI\_TaxID=1280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VARIOUS STRAINS;  
 RX MEDLINE=20187516; PubMed=10722640;  
 RA Cramton S.E., Schnell N.F., Gotz F., Bruckner R.;  
 RT "Identification of a new repetitive element in Staphylococcus  
 aureus.";  
 RL Infect. Immun. 68:2344-2348(2000).  
 DR EMBL; AF195967; AAF60251.1; -.  
 DR EMBL; AF195963; AAF60243.1; -.  
 DR EMBL; AF195964; AAF60245.1; -.  
 DR EMBL; AF195965; AAF60247.1; -.  
 DR EMBL; AF195966; AAF60249.1; -.  
 DR InterPro; IPR001189; -.  
 DR ProDom; PD000475; -; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 11 AA; 1262 MW; 4F978F86A8B1A723 CRC64;

Query Match 20.0%; Score 3; DB 2; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 8.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 EKL 7  
 |||  
 Db 3 EKL 5

RESULT 15  
 Q9VN99 PRELIMINARY; PRT; 11 AA.  
 AC Q9VN99;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE CG2676 PROTEIN.  
 GN CG2676.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Butsam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J., Ye J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Rao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003603; AAF32049.1; -.  
 DR FlyBase; FBgn0037309; CG2676.  
 SQ SEQUENCE 11 AA; 1113 MW; 8760FE9BD5A1B1B1 CRC64;

Query Match 20.0%; Score 3; DB 5; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 8.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSK 4  
 |||  
 Db 9 SSK 11

Search completed: September 15, 2001, 12:45:38  
 Job time: 148 sec



Query Match 20.0%; Score 3; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LRR 9  
 |||  
 DB 7 LRR 9

RESULT 10  
 Q51812 PRELIMINARY; PRT; 9 AA.  
 ID Q9GK05  
 AC Q9GK05  
 DT 01-MAR-2001 (TReMBLrel. 16, Created)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE ALPHA S2-CASEIN (FRAGMENT).  
 GN CSNIS2  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Veltri C., Pilla F., Lagonigro R.;  
 RT "A new allele of goat alpha s2-casein.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ297312; CAC21706.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 9  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1013 MW; 490FA5A1B1AEB723 CRC64;

Query Match 20.0%; Score 3; DB 6; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EKL 7  
 |||  
 DB 1 EKL 3

RESULT 11  
 Q51812 PRELIMINARY; PRT; 10 AA.  
 ID Q51812  
 AC Q51812  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE DNA HELICASE (FRAGMENT).  
 GN TRAI.  
 OS Escherichia coli.  
 OG Plasmid F.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cram D.S., Loh S.M., Cheah K.C.C., Skurray R.A.;  
 RX MEDLINE=92009201; PubMed=1916281;  
 RT "Sequence and conservation of genes at the distal end of the transfer  
 region on plasmids F and R6-5.";  
 RL Gene 104:85-90(1991).  
 DR EMBL: M38047; AAA98090.1; -;  
 KW Helicase; Plasmid.  
 FT NON\_TER 10  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1106 MW; 59A2417731A33B13 CRC64;

Query Match 20.0%; Score 3; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 7.6e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 KEK 6  
 |||  
 DB 3 KEK 5

RESULT 12  
 Q47345 PRELIMINARY; PRT; 11 AA.  
 ID Q47345  
 AC Q47345  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE LEADER PEPTIDE.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Faber F., van Giezen M., van Gorcom R.F.M., Harder W.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 2-11 FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=85134883; PubMed=6396419;  
 RA Hudson G.S., Davidson B.E.;  
 RT "Nucleotide sequence and transcription of the phenylalanine and  
 tyrosine operons of Escherichia coli K12.";  
 RL J. Mol. Biol. 180:1023-1051(1984).  
 DR EMBL: Z70523; CAA94435.1; -;  
 SQ SEQUENCE 11 AA; 1402 MW; 87AB199204141775 CRC64;

Query Match 20.0%; Score 3; DB 2; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 8.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RRE 10  
 |||  
 DB 6 RRE 8

RESULT 13  
 Q9R7U8 PRELIMINARY; PRT; 11 AA.  
 ID Q9R7U8  
 AC Q9R7U8  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE GENE FOR DNR PROTEIN (A REGULATORY PROTEIN FOR THE EXPRESSION OF THE  
 NIR AND NOR GENES), COMPLETE CDS (FRAGMENT).  
 OS Pseudomonas aeruginosa  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PA01;  
 RX MEDLINE=95226457; PubMed=7711073;  
 RA Arai H., Igarashi Y., Kodama T.;  
 RT "The structural genes for nitric oxide reductase from Pseudomonas  
 aeruginosa.";  
 RL Biochim. Biophys. Acta 1261:279-284(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PA01;  
 RX MEDLINE=95394152; PubMed=7664887;  
 RA Arai H., Igarashi Y., Kodama T.;  
 RT "Expression of the nir and nor genes for denitrification of  
 Pseudomonas aeruginosa requires a novel CRP/FNR-related

O18854  
 ID O18854 PRELIMINARY; PRT; 8 AA.  
 AC O18854;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)  
 DE CD19 ANTIGEN (FRAGMENT).  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Liu P.-C., Shibuya H., Katz M.L., Johnson G.S.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF024717; AAB1967.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 832 MW; 6735A1ADDB1325A7 CRC64;

Query Match 20.0%; Score 3; DB 6; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SKE 5  
 Db 2 SKE 4

RESULT 7  
 O95574  
 ID O95574 PRELIMINARY; PRT; 9 AA.  
 AC O95574;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE SERUM ALBUMIN (FRAGMENT).  
 GN ALB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=FETAL LIVER;  
 RX MEDLINE=82081882; PubMed=6171778;  
 RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,  
 RA Najarian R.C., Seeburg P.H., Wion K.L.;  
 RT "The sequence of human serum albumin cDNA and its expression in E.  
 RT coli.";  
 RL Nucleic Acids Res. 9:6103-6114(1981).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=FETAL LIVER;  
 RX MEDLINE=82105994; PubMed=6275391;  
 RA Dugaiczky A., Law S.W., Dennison O.E.;  
 RT "Nucleotide sequence and the encoded amino acids of human serum  
 RT albumin mRNA.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=FETAL LIVER;  
 RX MEDLINE=85153492; PubMed=6085063;  
 RA Urano Y., Sakai M., Watanabe K., Tamaoki T.;  
 RT "Tandem arrangement of the albumin and alpha-fetoprotein genes in the  
 RT human genome.";  
 RL Gene 32:255-261(1984).  
 DR EMBL; L00133; AAC95394.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1186 MW; 92340B0400440681 CRC64;

Query Match 20.0%; Score 3; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 RER 11  
 Db 6 RER 8

RESULT 8  
 Q9UE09  
 ID Q9UE09 PRELIMINARY; PRT; 9 AA.  
 AC Q9UE09;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 16, Last annotation update)  
 DE MYRISTOYL-COA:PROTEIN N-MYRISTOYLTRANSFERASE (FRAGMENT).  
 GN NMT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98343933; PubMed=9677304;  
 RA McIlhinney R.A., Young K., Egerton M., Camble R., White A.,  
 RA Soloviev M.;  
 RT "Characterization of human and rat brain myristoyl-CoA:protein N-  
 RT myristoyltransferase: evidence for an alternative splice variant of  
 RT the enzyme.";  
 RL Biochem. J. 333:491-495(1998).  
 DR EMBL; Y17208; CAA76685.1; -;  
 DR InterPro; IPR001039; -;  
 DR ProDom; PD000050; -; 1.  
 KW Transferase.  
 FT NON\_TER 9  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1030 MW; 097D0B1DC76735B4 CRC64;

Query Match 20.0%; Score 3; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 RER 12  
 Db 7 RER 9

RESULT 9  
 Q9UE26  
 ID Q9UE26 PRELIMINARY; PRT; 9 AA.  
 AC Q9UE26;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 16, Last annotation update)  
 DE PEG1/MEST PROTEIN (FRAGMENT).  
 GN PEG1/MEST.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97336048; PubMed=9192843;  
 RA Riesenwijk A.M., Hu L., Schulz U., Tariverdian G., Hoeglund P.,  
 RA Kere J., Ropers H.H., Kalscheuer V.M.;  
 RT "Monoallelic expression of human PEG1/MEST is paralleled by parent-  
 RT specific methylation in fetuses.";  
 RL Genomics 42:236-244(1997).  
 DR EMBL; Y10620; CAA71631.1; -;  
 FT NON\_TER 9  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 1258 MW; C52D07340AB41412 CRC64;

OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82080657; PubMed=7312036;  
 RA Hudson L., Rossi J., Landy A.;  
 RT "Dual function transcripts specifying tRNA and mRNA."  
 RL Nature 294:422-427(1981).  
 DR EMBL; X04181; CAA27777.1; -.  
 KW Elongation factor; Protein biosynthesis.  
 FT NON\_TER 13  
 SQ SEQUENCE 13 AA; 1617 MW; C433BE82A18D0B19 CRC64;

Query Match 26.7%; Score 4; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SKEK 6  
 Db 2 SKEK 5

## RESULT 3

Q920G5 ID Q920G5 PRELIMINARY; PRT; 14 AA.  
 AC Q920G5;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
 DE SHORTER ISOFORM OF INTERLEUKIN 15 (FRAGMENT).  
 GN INTERLEUKIN 15.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALBC;  
 RX Nishimura H., Washizu J., Nakamura N., Enomoto A., Yoshikai Y.;  
 RT "Translational efficiency is up-regulated by alternative exon in murine IL-15 mRNA."  
 RL J. Immunol. 160:936-942(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Prinz M., Hanisch U.K., Kettenmann H., Kirchhoff F.;  
 RT "Alternative splicing of mouse IL-15 is due to the use of an internal splice site in exon 5."  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB022307; BAA37122.1; -.  
 DR EMBL; AJ012587; CAA10069.1; -.  
 FT NON\_TER 14  
 SQ SEQUENCE 14 AA; 1503 MW; 62A43C3889924E2B CRC64;

Query Match 26.7%; Score 4; DB 11; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSKE 5  
 Db 4 SSKE 7

## RESULT 4

Q66205 ID Q66205 PRELIMINARY; PRT; 7 AA.  
 AC Q66205;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
 DE MEMBRANE PROTEIN (1 IS 3RD BASE IN CODON) (FRAGMENT).  
 OS porcine transmissible gastroenteritis virus.  
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 CC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=11149;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FS772/70;  
 RX MEDLINE=88216185; PubMed=2835592;  
 RA Britton P., Carmenes R.S., Page K.W., Garwes D.J., Parra F.;  
 RT "Sequence of the nucleoprotein gene from a virulent British field isolate of transmissible gastroenteritis virus and its expression in Saccharomyces cerevisiae."  
 RL Mol. Microbiol. 2:89-99(1988).  
 DR EMBL; Y00542; CAA68606.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 869 MW; 72C691E727233B70 CRC64;

Query Match 20.0%; Score 3; DB 14; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 EKL 7  
 Db 1 EKL 3

## RESULT 5

Q9N6M5 ID Q9N6M5 PRELIMINARY; PRT; 8 AA.  
 AC Q9N6M5;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
 DE DITHYDROFOLATE REDUCTASE THYMIDYLATE SYNTHASE (FRAGMENT).  
 GN FOLL.  
 OS Toxoplasma gondii.  
 CC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;  
 CC Toxoplasma.  
 OX NCBI\_TaxID=5811;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RH, COUGAR TC751G34, SEA OTTER TC828G1, AND BEVERLEY;  
 RA Lehmann T., Blackston C.R., Farmlay S.F., Remington J.S., Dubey J.P.;  
 RT "Strain Typing of Toxoplasma gondii: Comparison of Antigen-Coding and Housekeeping Genes."  
 RL J. Parasitol. 0:0-0(2000).  
 DR EMBL; AF249695; AAF79153.1; -.  
 DR EMBL; AF249692; AAF79150.1; -.  
 DR EMBL; AF249693; AAF79151.1; -.  
 DR EMBL; AF249694; AAF79152.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 1035 MW; 33CAAAA05B133044 CRC64;

Query Match 20.0%; Score 3; DB 5; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 RIK 13  
 Db 1 RIK 3

## RESULT 6

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2001, 12:43:10 ; Search time 17.55 Seconds

(without alignments)  
113.081 Million cell updates/sec

Title: US-09-389-000-2\_COPY\_140\_154

Perfect score: 15

Sequence: 1 HSSKEKLRRIKYC 15

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2508

Minimum DB seq length: 0

Maximum DB seq length: 15

Post-processing: Listing first 45 summaries

- Database : SPTREMBL\_16:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_unclassified:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	26.7	11	10 Q9S8X4	Q9S8X4 glycine max
2	4	26.7	13	2 Q47693	Q47693 escherichia
3	4	26.7	14	11 Q9Z0G5	Q9Z0G5 mus musculus
4	3	20.0	7	14 Q66205	Q66205 porcine tra
5	3	20.0	8	5 Q9N6M5	Q9N6M5 toxoplasma
6	3	20.0	8	6 O18854	O18854 canis famil
7	3	20.0	9	4 Q95574	Q95574 homo sapien
8	3	20.0	9	4 Q9UE09	Q9UE09 homo sapien
9	3	20.0	9	4 Q9UE26	Q9UE26 homo sapien
10	3	20.0	9	6 Q9GK05	Q9GK05 capra hircu
11	3	20.0	10	2 O51812	O51812 escherichia
12	3	20.0	11	2 Q47345	Q47345 escherichia
13	3	20.0	11	2 Q9R708	Q9R708 pseudomonas
14	3	20.0	11	2 Q9K332	Q9K332 staphylococ
15	3	20.0	11	5 Q9VW99	Q9VW99 drosophila
16	3	20.0	11	10 Q04131	Q04131 lycopersico
17	3	20.0	12	2 O50303	O50303 bacillus st
18	3	20.0	12	4 Q9UMR0	Q9UMR0 homo sapien
19	3	20.0	12	4 Q9UMQ9	Q9UMQ9 homo sapien

20	3	20.0	12	6 Q9XT42	Q9XT42 canis famil
21	3	20.0	13	2 Q55234	Q55234 synechocyst
22	3	20.0	13	4 Q9UM46	Q9UM46 homo sapien
23	3	20.0	13	5 Q9W506	Q9W506 drosophila
24	3	20.0	13	7 Q29823	Q29823 homo sapien
25	3	20.0	14	2 Q47335	Q47335 escherichia
26	3	20.0	14	4 O15222	O15222 homo sapien
27	3	20.0	14	10 Q9XGM4	Q9XGM4 arabidopsis
28	3	20.0	14	10 Q9FXW3	Q9FXW3 nicotiana s
29	3	20.0	14	11 Q9QW75	Q9QW75 mus sp. hom
30	3	20.0	14	14 O10234	O10234 human immun
31	3	20.0	15	1 Q9UWG9	Q9UWG9 thermococcu
32	3	20.0	15	2 O52059	O52059 salmonella
33	3	20.0	15	4 P78482	P78482 homo sapien
34	3	20.0	15	4 Q9UMT6	Q9UMT6 homo sapien
35	3	20.0	15	4 Q9UBK0	Q9UBK0 homo sapien
36	3	20.0	15	4 Q9UCES	Q9UCES homo sapien
37	3	20.0	15	4 Q9UC83	Q9UC83 homo sapien
38	3	20.0	15	10 Q42223	Q42223 arabidopsis
39	3	20.0	15	11 O35411	O35411 mus musculu
40	2	13.3	6	2 P82655	P82655 lactobacill
41	2	13.3	7	2 P70804	P70804 azotobacter
42	2	13.3	7	4 Q15903	Q15903 homo sapien
43	2	13.3	7	8 P92421	P92421 psathyrosta
44	2	13.3	7	8 P92385	P92385 hordeum mar
45	2	13.3	7	8 P92210	P92210 agropyron c

ALIGNMENTS

RESULT 1

Q9S8X4 ID Q9S8X4 PRELIMINARY; PRT; 11 AA.  
AC Q9S8X4  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE VEGETATIVE STORAGE PROTEIN 94 PEPTIDE 3, VSP94-LIPOXYGENASE.  
OS Glycine max (Soybean).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
OC Fabales; Fabaceae; Papilionoideae; Glycine.  
OX NCBI\_TaxID=3847;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92361246; PubMed=1822994;  
RA Tranbarger T.J., Franceschi V.R., Hildebrand D.F., Grimes H.D.;  
RT "The soybean 94-kilodalton vegetative storage protein is a  
RT lipoxigenase that is localized in paraveinal mesophyll cell  
RT vacuoles."  
RL Plant Cell 3:973-987(1991).  
SQ SEQUENCE 11 AA; 1366 MW; 9B337C3C0DD9CB1A CRC64;

Query Match 26.7%; Score 4; DB 10; Length 11;  
Best Local Similarity 100.0%; Pred. No. 9.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 RERI 12  
Db 8 RERI 11

RESULT 2

Q47693 ID Q47693 PRELIMINARY; PRT; 13 AA.  
AC Q47693  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE ELONGATION FACTOR TU (EF-TU) (FRAGMENT).  
GN TUFB.



```

/lab_host="DH12S"
/Note="Organ: whole embryo; Vector: pCMV-SPORT; Site_1:
Salt; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dt. Gastrulating embryos were collected at 7.5dpc
from C57Bl6 x DBA matings, excluding embryos that had
developed head folds and all extraembryonic tissues.
Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
Referenced in Development 121, 2479-2489 (1995)."
BASE COUNT      69 a   47 c   46 g   87 t
ORIGIN

```

```

Query Match      1.6%; Score 20; DB 3; Length 249;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 72 aaagaaataacacattca 91
|||||
Db 191 AAAGAAAATACACATTCA 172

```

```

RESULT 15
BE606386/c
LOCUS      BE606386      319 bp      mRNA      EST      22-AUG-2000
DEFINITION WHE0904_D06_H12S Wheat 5-15 DAP spike cDNA library Triticum
aestivum CDNA clone WHE0904_D06_H12, mRNA sequence.
ACCESSION BE606386
VERSION    BE606386.1 GI:9883550
KEYWORDS   EST.
SOURCE     bread wheat.
ORGANISM   Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.

```

```

REFERENCE
AUTHORS    Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
TITLE      The structure and function of the expressed portion of the wheat
genomes - 5-15 DAP spike cDNA library
JOURNAL    Unpublished (2000)
COMMENT    Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595618
Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

```

```

FEATURES
source
1..319
Location/Qualifiers
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE0904_D06_H12"
/clone_lib="Wheat 5-15 DAP spike cDNA library"
/tissue_type="Spike"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/Note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
greenhouse. Spikes at 5, 10 and 15 DAP were harvested,
total RNA and poly(A) RNA were prepared, a cDNA library
was made, and the cDNA clones were in vivo excised to
give Bluescript phagemids in the TJ Clonase lab (Choi,
Close, Fenton) at the University of California,
Riverside. Plasmid DNA preparations and DNA sequencing
were performed in the OD Anderson lab (all other authors
)."

```

```

BASE COUNT      78 a   86 c   108 g   47 t
ORIGIN

```

```

Query Match      1.6%; Score 20; DB 137; Length 319;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 981 ccccggtctcaagcttctcc 1000
|||||
Db 111 CCCCgGTCTCAAGCTTCTCC 92

```

```

Search completed: September 15, 2001, 10:55:41
Job time: 5044 sec

```

JOURNAL COMMENT Unpublished (2000)  
Contact: Cathy Ronning  
The Institute for Genomic Research  
For clone info: please contact Research Genetics, Libraries  
Division tel 1-800-711-6195, email cdna@resgen.com  
Seq primer: M13P-R.

FEATURES source Location/Qualifiers  
1. .606  
/organism="Solanum tuberosum"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="BPL1213"  
/clone\_lib="P. infestans-challenged leaf"  
/tissue\_type="leaf"  
/dev\_stage="6 week old"  
/lab\_host="SOLR"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; Whole plants were challenged with 450,000 sporangia/ml P. infestans US-1(US 940501) in Biotron (Madison, Wisconsin). Leaf tissue was collected at 1, 2, 5, 12, and 24 hours post-challenge and frozen in liquid nitrogen immediately upon removal. Kennebec plants showed no signs of HR. Katahdin plants (susceptible to P. infestans US-1) were used as controls and showed infection. NOTE: We cannot exclude the possibility that this sequence is actually derived from Phytophthora rather than potato."

BASE COUNT 179 a 100 c 141 g 186 t  
ORIGIN

Query Match 1.7%; Score 21; DB 155; Length 606;  
Best Local Similarity 100.0%; Pred. No. 6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 402 agatctctcttcattca 422  
|||||  
Db 4 AGATCTCTTCTTCATCAA 24

RESULT 13  
LOCUS AQ659253/c  
DEFINITION Sheared DNA-3D4-TR Sheared DNA Trypanosoma brucei genomic clone  
Sheared DNA-3D4, DNA sequence.  
ACCESSION AQ659253  
VERSION AQ659253.1 GI:5167021  
KEYWORDS GSS.  
SOURCE Trypanosoma brucei.  
ORGANISM Trypanosoma brucei  
Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;  
Trypanosoma.  
REFERENCE 1 (bases 1 to 625)  
AUTHORS El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C., Gerrard,C., Leech,V.; de Jong,P., Ullu,E., Melville,S., Donelson,J., Fraser,C. and Adams,M.  
Determination of clone end sequences from Trypanosoma brucei GUTat 10.1 sheared DNA library  
Unpublished (1999)  
Other\_GSSs: Sheared DNA-3D4.TF  
Contact: Najib M. El-Sayed  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: nelsayed@tigr.org  
Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared DNA library constructed at TIGR. Clones will be available for distribution through ATCC. Sheared DNA end sequences search page: <http://www.tigr.org/tdb/mdb/tbdb/>.  
Seq primer: M13-Reverse  
Class: shotgun.

FEATURES source Location/Qualifiers  
1. .625  
/organism="Trypanosoma brucei"  
/strain="TREU927/4 GUTat 10.1"  
/db\_xref="taxon:5691"  
/clone="Sheared DNA-3D4"  
/clone\_lib="Sheared DNA"  
/note="Vector: pUC18; Site\_1: SmaI; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (approx 2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Borell, Oxford University Press, 1999)."

BASE COUNT 174 a 155 c 154 g 142 t  
ORIGIN

Query Match 1.7%; Score 21; DB 231; Length 625;  
Best Local Similarity 100.0%; Pred. No. 6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 488 tatgtaaaaggagagaagaat 508  
|||||  
Db 115 TATGTAAAAGGGAGAGAAT 95

RESULT 14  
LOCUS AA155374/c  
DEFINITION aa43hl2.r1 Beddington mouse embryonic region Mus musculus cDNA  
clone IMAGE:540743 5', mRNA sequence.  
ACCESSION AA155374  
VERSION AA155374.1 GI:1727004  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 249)  
AUTHORS Marra,M., Hallier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.  
The WashU-HHMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
MG1:327679  
Seq primer: -40ml3 ET  
High quality sequence stop: 236.  
Location/Qualifiers  
1. .249  
/organism="Mus musculus"  
/strain="C57BL6 x DBA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:540743"  
/clone\_lib="Beddington mouse embryonic region"  
/sex="pooled"  
/tissue\_type="embryo"  
/dev\_stage="7.5dpc"

```

BASE COUNT      45 a 49 c 5 g 71 t
ORIGIN

Query Match      1.7%; Score 21; DB 166; Length 170;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 agatctcttcttcttcaaa 422
|||||
Db 109 AGATCTCTCTCTTCATTCAA 129

RESULT 10
BF187644
LOCUS      BF187644      471 bp      mRNA      EST      01-NOV-2000
DEFINITION EST443931 potato stolon, Cornell University Solanum tuberosum cDNA
clone cSTA41B22 5' sequence, mRNA sequence.
ACCESSION  BF187644
VERSION     BF187644.1 GI:11069863
KEYWORDS    EST.
SOURCE      potato.
ORGANISM    Solanum tuberosum
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1 (bases 1 to 471)
AUTHORS    van der Hoeven,R., Bezzerides,J., Bachem,C., Horvath,B., Visser,R.,
Holt,I.E., Liang,F., Hansen,T.S., Utterback,T., Bowman,C.L., Doan
,B., Bougri,O., Buell,C.R., Ronning,C.M., Tanksley,S.D. and Baker
,B.
TITLE      Generation of ESTs from potato swelling stolons
JOURNAL
COMMENT    Unpublished (1999)
Contact: TIGR
The Institute for Genomic Research
Rockville, MD 20850, USA
Tel: (301)838-0200
Email: http://www.tigr.org/tdb/potato/
For clone information, please contact Research Genetics, Libraries
Division, tel# 1-800-711-6195, email- cdna@resgen.com.

FEATURES             source
1..471
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="cSTA41B22"
/tissue_type="axillary buds of stem explants, swelling
stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; RNA was supplied by Christian Bachem & Beatrix
Horvath(Laboratory of Plant Breeding, Dept. of Plant
Sciences, Wageningen University, The Netherlands). Total
RNA was isolated from developing axillary buds of potato
nodal stem cuttings cultured on medium for the
introduction of tuber formation as described in Bachem et
al. (Plant Journal 1996). Tissue samples were taken of
stages corresponding to growing stolons and the early
stages of tuber formation."
BASE COUNT      125 a 110 c 68 g 168 t
ORIGIN

Query Match      1.7%; Score 21; DB 145; Length 471;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 agatctcttcttcttcaaa 422
|||||
Db 175 AGATCTCTCTCTTCATTCAA 195

us-09-389-000-1_copy_735_1949.rst
Page 7

```



```

and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCCTCAGTCAGCAG
Plate: 40 row: L column: 7
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
source
1. 553
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/Note="vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 163 a 106 c 121 g 163 t
ORIGIN

Query Match 2.1%; Score 26; DB 144; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 116 aattttaaaggtgtatttcaggcca 141
|||||
Db 247 AATTTAAAGGTTGATTTCAGGGCA 272

RESULT 9
BE342634 170 bp mRNA EST 17-JUL-2000
LOCUS
DEFINITION BE342634 potato stolon, Cornell University Solanum tuberosum cDNA
clone cSTA20P1, mRNA sequence.
ACCESSION BE342634
VERSION BE342634.1 GI:9252166
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 170)
AUTHORS van der Hoeven, R., Bezzerides, J., Bachem, C., Horvath, B., Visser, R.,
Holt, I.E., Liang, F., Hansen, T.S., Utterback, T., Bowman, C.L., Doan
, B., Bougri, O., Buell, C.R., Ronning, C.M., Tanksley, S.D. and Baker
, B.
Generation of ESTs from potato swelling stolons
Unpublished (1999)
Contact: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cdna@resgen.com
5 prime sequence.
Location/Qualifiers
source
1. 170
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="cSTA20P1"
/clone_lib="potato stolon, Cornell University"
/tissue_type="axillary buds of stem explants, swelling
stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/Note="vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; RNA was supplied by Christian Bachem & Beatrix
Horvath(Laboratory of Plant Breeding, Dept. of Plant
Sciences, Wageningen University, The Netherlands). Total
RNA was isolated from developing axillary buds of potato
nodal stem cuttings cultured on medium for the
introduction of tuber formation as described in Bachem et
al. (Plant Journal 1996). Tissue samples were taken of
stages corresponding to growing stolons and the early

```

---

```

1
RESULT 7
AW961078 624 bp mRNA EST 01-JUN-2000
LOCUS
DEFINITION ES373045 MAGE resequences, MAGF Homo sapiens cDNA, mRNA sequence.
ACCESSION AW961078
VERSION AW961078.1 GI:8150657
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 624)
AUTHORS Hegde, P., Qi, R., Athernath, K., Dharap, S., Gaspard, R., Gay, C., Holt
, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
Quackenbush, J.
Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
Unpublished (2000)
Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 152
Seq primer: Forward.
Location/Qualifiers
source
1. 624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGF"
/Note="vector: pBluescriptSKm"
BASE COUNT 209 a 117 c 128 g 170 t
ORIGIN

Query Match 3.5%; Score 43; DB 122; Length 624;
Best Local Similarity 100.0%; Pred. No. 5.1e-12;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 ttcaatagtcgagagctggaagccatcaagtaattaga 64
|||||
Db 462 TTCATAAGTCGAGGAGCTGGAGCCATCAAGTAATTAGA 504

RESULT 8
BF080785 553 bp mRNA EST 18-OCT-2000
LOCUS
DEFINITION 231799 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF080785
VERSION BF080785.1 GI:10874615
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 553)
AUTHORS Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.
and Keele, J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18

```

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html).

Seq primer: M13-21

Class: BAC ends.

#### FEATURES

source

1..452

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="2383N20"

/clone\_lib="CIT-HSP"

/sex="Male"

/cell\_type="Sperm"

/note="vector: pBelOBAC11; Site\_1: HindIII; Site\_2: HindIII"

BASE COUNT 128 a 88 c 88 g 148 t

ORIGIN

Query Match 7.8%; Score 95; DB 225; Length 452;  
Best Local Similarity 100.0%; Pred. No. 1.4e-40;  
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 588 agattacagaagcattcagagcaacatgaggtttgtaagaaacaaacacccattg 647

|||||

Db 189 AGATTACAGAAGCACTTCAGAGCAACATGAGGTTTGTAAAGAAACAACAAACCCATTG 248

|||||

Qy 648 agctctctccagcactgtcatggcacagcg 682

|||||

Db 249 ACCTGTCCTCCAGCACTGTCATGGCACACGG 283

RESULT 5  
LOCUS AQ267168 438 bp DNA GSS 27-APR-1999  
DEFINITION RPC111-73A7.TJ RPC1-11 Homo sapiens genomic clone RPC1-11-73A7, DNA sequence.  
ACCESSION AQ267168  
VERSION AQ267168.1 GI:3794772  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 438)  
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.

Use of human BAC End Sequences for Sequence-Ready Map Building

Other\_GSSs: RPC111-73A7.TK

Unpublished (1998)

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are derived from the human BAC library RPC1-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from

Research Genetics (info@resgen.com). BAC end search page:

[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html)

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1..438

/organism="Homo sapiens"

/db\_xref="GDB:7527654"

FEATURES

source

/db\_xref="taxon:9606"

/clone="RPC1-11-73A7"

/clone\_lib="RPC1-11"

/sex="Male"

/cell\_type="Lymphocytes"

/note="vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;

RPC111 Human Male BAC Library"

BASE COUNT 153 a 81 c 78 g 126 t

ORIGIN

Query Match 4.0%; Score 48; DB 226; Length 438;  
Best Local Similarity 100.0%; Pred. No. 9e-15;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 799 agttccatcaagctccgctcagagaatgctattggtatccatataa 846

|||||

Db 48 AGTTCATCAAGCTCCGCTCAGAGATGCTATTGTTGATCCATATAA 1

|||||

RESULT 6

B92720

LOCUS B92720 354 bp DNA GSS 25-JUN-1998

DEFINITION CIT-HSP-2170M15.TF CIT-HSP Homo sapiens genomic clone 2170M15, DNA sequence.

ACCESSION B92720

VERSION B92720.1 GI:2975057

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 354)

Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,

Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,

Simon,M. and Venter,J.C.

Use of a random BAC End Sequence Database for Sequence-Ready Map

Building (1998)

Unpublished (1998)

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html)

Seq primer: M13-21;

Class: BAC ends.

Location/Qualifiers

1..354

/organism="Homo sapiens"

/db\_xref="GDB:7103209"

/db\_xref="taxon:9606"

/clone="2170M15"

/clone\_lib="CIT-HSP"

/sex="Male"

/cell\_type="Sperm"

/note="vector: pBelOBAC11; Site\_1: HindIII; Site\_2: HindIII"

BASE COUNT 93 a 70 c 67 g 124 t

ORIGIN

Query Match 3.8%; Score 46; DB 257; Length 354;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 588 agattacagaagcattcagagcaacatgaggtttgtaagaaaca 633

|||||

Db 200 AGATTACAGAAGCACTTCAGAGCAACATGAGGTTTGTAAAGAAACA 245

|||||

```

VERSION      AA417643.1  GI:2079462
KEYWORDS     EST.
SOURCE       HOMO sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE    Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS      Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
              Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
              , J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
              White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE        WashU-NCI human EST Project
JOURNAL      Unpublished (1997)
COMMENT      Contact: Wilton RK
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              This clone is available royalty-free through LNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              Insert Length: 950 Std Error: 0.00
              Seq primer: -28ml3 rev2 ET from Amersham
              High quality sequence stop: 465.

FEATURES     source
              Location/Qualifiers
                1..490
                /organism="Homo sapiens"
                /db_xref="GDB:5975465"
                /db_xref="taxon:9606"
                /clone="IMAGE:752630"
                /clone_lib="Soares-NHMPu_S1"
                /tissue_type="pooled human melanocyte, fetal heart, and
                pregnant uterus"
                /lab_host="DH10B"
                /note="Organ: mixed (see below); Vector: pWT73D-Pac
                (Pharmacia) with a modified polylinker; Site_1: Not I;
                Site_2: Eco RI; Equal amounts of plasmid DNA from three
                normalized libraries (melanocyte 2NBHM, pregnant uterus
                NBHPU, and fetal heart NBHH19W) were mixed, and ss circles
                were made in vitro. Following HAP purification, this DNA
                was used as tracer in a subtractive hybridization
                reaction. The driver was PCR-amplified cDNAs from pools of
                5,000 clones made from the same 3 libraries. The pools
                consisted of I.M.A.G.E. clones 260232-265223,
                340488-345479, and 484488-489479."
BASE COUNT   158 a 114 c 97 g 121 t
ORIGIN

Query Match      15.6%; Score 189; DB 7; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.8e-92;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1027 cagacgtcacatcacatcccaactgtctgcaacagtttttggcggtattaaagcacatg 1086
      |||||||
Db 35 CAGAGCTGCATACATACATCCCAACTGTCTGCAACAGTTTGGGGCGTATTAAGACCATG 94
      |||||||
Qy 1087 ttggaattcacactctcaacacactactggcgcgagtttgacaatctaggaaaagtga 1146
      |||||||
Db 95 TTTGAATTCACACTCTCAACCCACTACTGGCGCAGTTTCACAATCTAGGAAAAGTGG 154
      |||||||
Qy 1147 acaagaatgattttgaagtcaccccaagacctaataatcaaaagagttggcattggt 1206
      |||||||
Db 155 ACAAGAATGATTTTGAAGACTCCACCCAAAGACCTAATATCAAAAGAGTTGGCATGTT 214
      |||||||
Qy 1207 tggctcttg 1215
      |||||||
Db 215 TGGCTCTG 223

RESULT      3
AW961069
LOCUS

Query Match      15.6%; Score 189; DB 7; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.8e-92;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1027 cagacgtcacatcacatcccaactgtctgcaacagtttttggcggtattaaagcacatg 1086
      |||||||
Db 35 CAGAGCTGCATACATACATCCCAACTGTCTGCAACAGTTTGGGGCGTATTAAGACCATG 94
      |||||||
Qy 1087 ttggaattcacactctcaacacactactggcgcgagtttgacaatctaggaaaagtga 1146
      |||||||
Db 95 TTTGAATTCACACTCTCAACCCACTACTGGCGCAGTTTCACAATCTAGGAAAAGTGG 154
      |||||||
Qy 1147 acaagaatgattttgaagtcaccccaagacctaataatcaaaagagttggcattggt 1206
      |||||||
Db 155 ACAAGAATGATTTTGAAGACTCCACCCAAAGACCTAATATCAAAAGAGTTGGCATGTT 214
      |||||||
Qy 1207 tggctcttg 1215
      |||||||
Db 215 TGGCTCTG 223

DEFINITION    EST373036 MAGE resequences, MAGF Homo sapiens cDNA, mRNA sequence.
ACCESSION     AW961069
VERSION       AW961069.1  GI:8150648
KEYWORDS      EST.
SOURCE        HOMO sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE     Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS       Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt
              , I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
              Quackenbush, J.
TITLE         Assessment of gene expression patterns in a model of colon tumor
              metastasis using a 19,200 element cDNA microarray
JOURNAL       Unpublished (2000)
COMMENT       Contact: John Quackenbush
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 3528
              Fax: 301 838 0208
              Email: johnq@tigr.org
              Plate: 152
              Seq primer: Forward.
              Location/Qualifiers
                1..617
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="MAGE resequences, MAGF"
                /note="vector: pBluescriptskm"
BASE COUNT    211 a 116 c 127 g 163 t
ORIGIN

Query Match      10.4%; Score 126; DB 122; Length 617;
Best Local Similarity 99.4%; Pred. No. 1.4e-57;
Matches 176; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 9gtctcttggaagtgcccttctcactaagtgcgcgagagctggaagccatcaagtaatt 60
      |||||||
Db 441 GGTCTCTTGAAGTGCCCTCTTCTCACTAAGTCCGAGGAGCTGGAAGCCATCAAGTTAAT 500
      |||||||
Qy 61 tagattggcaaaaaagaataacacattcactgtttgtttttataatccctgaaattt 120
      |||||||
Db 501 TAGATTGGCAAAAAAGAATAACACATTCATCTGTTCTTTTATAATCCCTGAAATTT 560
      |||||||
Qy 121 taaaggttgatttcagggcgatggaatgatattgcttttaactgaaccactgacaat 177
      |||||||
Db 561 TAAAGTTGTATTTTCAGGGCATGGAATGATATTCCTTTAACTGAACCACTGACAA 617
      |||||||

DEFINITION    CIT-HSP-2383N20.TF CIT-HSP Homo sapiens genomic clone 2383N20, DNA
ACCESSION     A0196491
VERSION       A0196491
KEYWORDS      GSS.
SOURCE        human.
ORGANISM      Homo sapiens
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
              1 (bases 1 to 452)
              Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
              Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and
              Venter, J.C.
TITLE         Use of a random human BAC End Sequence Database for Sequence-Ready
              Map Building
JOURNAL       Unpublished (1998)
COMMENT       Other GSSs: CIT-HSP-2383N20.TR
              Contact: Mark Adams
              Department of Eukaryotic Genomics
              The Institute for Genomic Research

```



117: gb\_est48:\*  
118: gb\_est49:\*  
119: gb\_est50:\*  
120: gb\_est51:\*  
121: gb\_est52:\*  
122: gb\_est53:\*  
123: gb\_est54:\*  
124: gb\_est55:\*  
125: gb\_est56:\*  
126: gb\_est57:\*  
127: gb\_est58:\*  
128: gb\_est59:\*  
129: gb\_est60:\*  
130: gb\_est61:\*  
131: gb\_est62:\*  
132: gb\_est63:\*  
133: gb\_est64:\*  
134: gb\_est65:\*  
135: gb\_est66:\*  
136: gb\_est67:\*  
137: gb\_est68:\*  
138: gb\_est69:\*  
139: gb\_est70:\*  
140: gb\_est71:\*  
141: gb\_est72:\*  
142: gb\_est73:\*  
143: gb\_est74:\*  
144: gb\_est75:\*  
145: gb\_est76:\*  
146: gb\_est77:\*  
147: gb\_est78:\*  
148: gb\_est79:\*  
149: gb\_est80:\*  
150: gb\_est81:\*  
151: gb\_est82:\*  
152: gb\_est83:\*  
153: gb\_est84:\*  
154: gb\_est85:\*  
155: gb\_est86:\*  
156: gb\_est87:\*  
157: gb\_est88:\*  
158: gb\_est89:\*  
159: gb\_est90:\*  
160: gb\_est91:\*  
161: gb\_est92:\*  
162: gb\_est93:\*  
163: gb\_est94:\*  
164: gb\_est95:\*  
165: gb\_est96:\*  
166: gb\_est97:\*  
167: gb\_est98:\*  
168: gb\_est99:\*  
169: gb\_est100:\*  
170: gb\_est101:\*  
171: gb\_est102:\*  
172: gb\_est103:\*  
173: gb\_est104:\*  
174: gb\_est105:\*  
175: gb\_est106:\*  
176: gb\_est107:\*  
177: gb\_est108:\*  
178: gb\_est109:\*  
179: gb\_est110:\*  
180: gb\_est111:\*  
181: gb\_est112:\*  
182: gb\_est113:\*  
183: gb\_est114:\*  
184: gb\_est115:\*  
185: gb\_est116:\*  
186: gb\_est117:\*  
187: gb\_est118:\*  
188: gb\_est119:\*  
189: gb\_est120:\*

190: gb\_est121:\*  
191: gb\_est122:\*  
192: gb\_est123:\*  
193: gb\_est124:\*  
194: gb\_est125:\*  
195: gb\_est126:\*  
196: gb\_est127:\*  
197: gb\_est128:\*  
198: gb\_est129:\*  
199: gb\_est130:\*  
200: gb\_est131:\*  
201: gb\_est132:\*  
202: gb\_est133:\*  
203: gb\_est134:\*  
204: gb\_est135:\*  
205: gb\_est136:\*  
206: gb\_est137:\*  
207: gb\_est138:\*  
208: gb\_est139:\*  
209: gb\_est140:\*  
210: gb\_est141:\*  
211: gb\_est142:\*  
212: gb\_est143:\*  
213: gb\_est144:\*  
214: gb\_est145:\*  
215: gb\_est146:\*  
216: gb\_est147:\*  
217: gb\_est148:\*  
218: gb\_est149:\*  
219: gb\_est150:\*  
220: gb\_est151:\*  
221: gb\_est152:\*  
222: gb\_est153:\*  
223: gb\_est154:\*  
224: gb\_est155:\*  
225: gb\_est156:\*  
226: gb\_est157:\*  
227: gb\_est158:\*  
228: gb\_est159:\*  
229: gb\_est160:\*  
230: gb\_est161:\*  
231: gb\_est162:\*  
232: gb\_est163:\*  
233: gb\_est164:\*  
234: gb\_est165:\*  
235: gb\_est166:\*  
236: gb\_est167:\*  
237: gb\_est168:\*  
238: gb\_est169:\*  
239: gb\_est170:\*  
240: gb\_est171:\*  
241: gb\_est172:\*  
242: gb\_est173:\*  
243: gb\_est174:\*  
244: gb\_est175:\*  
245: gb\_est176:\*  
246: gb\_est177:\*  
247: gb\_est178:\*  
248: gb\_est179:\*  
249: gb\_est180:\*  
250: gb\_est181:\*  
251: gb\_est182:\*  
252: gb\_est183:\*  
253: gb\_est184:\*  
254: gb\_est185:\*  
255: gb\_est186:\*  
256: gb\_est187:\*  
257: gb\_est188:\*  
258: gb\_est189:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2001, 09:31:37 ; Search time 1305.86 Seconds  
(without alignments)  
8795.137 Million cell updates/sec

Title: US-09-389-000-1\_COPY\_735\_1949  
Perfect score: 1215  
Sequence: 1 ggtctctgaagggcctt.....ttggcatggttggtctctg 1215

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 10228115 seqs, 4726426750 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST.\*

1: gb\_estl.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_est9.\*  
10: gb\_est10.\*  
11: gb\_est11.\*  
12: gb\_est12.\*  
13: gb\_est13.\*  
14: gb\_est14.\*  
15: gb\_est15.\*  
16: gb\_est16.\*  
17: gb\_est17.\*  
18: gb\_est18.\*  
19: gb\_est19.\*  
20: gb\_est20.\*  
21: gb\_est21.\*  
22: gb\_est22.\*  
23: gb\_est23.\*  
24: gb\_est24.\*  
25: gb\_est33.\*  
26: gb\_est34.\*  
27: gb\_est35.\*  
28: gb\_est36.\*  
29: gb\_est37.\*  
30: gb\_est38.\*  
31: gb\_est39.\*  
32: gb\_est40.\*  
33: em\_estba.\*  
34: em\_estfun.\*  
35: em\_esthum1.\*  
36: em\_esthum2.\*  
37: em\_esthum3.\*  
38: em\_esthum4.\*  
39: em\_esthum5.\*  
40: em\_esthum6.\*  
41: em\_esthum7.\*  
42: em\_esthum8.\*  
43: em\_esthum9.\*

44: em\_esthum10.\*  
45: em\_esthum11.\*  
46: em\_esthum12.\*  
47: em\_esthum13.\*  
48: em\_esthum14.\*  
49: em\_esthum15.\*  
50: em\_esthum16.\*  
51: em\_esthum17.\*  
52: em\_esthum18.\*  
53: em\_esthum19.\*  
54: em\_esthum20.\*  
55: em\_esthum21.\*  
56: em\_esthum22.\*  
57: em\_esthum23.\*  
58: em\_esthum24.\*  
59: em\_esthum25.\*  
60: em\_esthum26.\*  
61: em\_esthum27.\*  
62: em\_esthum28.\*  
63: em\_estin1.\*  
64: em\_estin2.\*  
65: em\_estin3.\*  
66: em\_estin4.\*  
67: em\_estin5.\*  
68: em\_estom1.\*  
69: em\_estom2.\*  
70: em\_estov1.\*  
71: em\_estov2.\*  
72: em\_estpl1.\*  
73: em\_estpl2.\*  
74: em\_estpl3.\*  
75: em\_estpl4.\*  
76: em\_estpl5.\*  
77: em\_estpl6.\*  
78: em\_estpl7.\*  
79: em\_estpl8.\*  
80: em\_estpl9.\*  
81: em\_estpl10.\*  
82: em\_estro1.\*  
83: em\_estro2.\*  
84: em\_estro3.\*  
85: em\_estro4.\*  
86: em\_estro5.\*  
87: em\_estro6.\*  
88: em\_estro7.\*  
89: em\_estro8.\*  
90: em\_estro9.\*  
91: em\_estro10.\*  
92: em\_estro11.\*  
93: em\_estro12.\*  
94: em\_estro13.\*  
95: em\_estro14.\*  
96: em\_estro15.\*  
97: em\_estro16.\*  
98: em\_estro17.\*  
99: em\_estro18.\*  
100: em\_estro19.\*  
101: em\_estro20.\*  
102: gb\_est25.\*  
103: gb\_est26.\*  
104: gb\_est27.\*  
105: gb\_est28.\*  
106: gb\_est29.\*  
107: gb\_est30.\*  
108: gb\_est31.\*  
109: gb\_est32.\*  
110: gb\_est41.\*  
111: gb\_est42.\*  
112: gb\_est43.\*  
113: gb\_est44.\*  
114: gb\_est45.\*  
115: gb\_est46.\*  
116: gb\_est47.\*

**THIS PAGE BLANK (USPTO)**



Conservative 0; Mismatches 0; Indels 0; Gaps 0;

cttcactaagtgc 33  
 |||||  
 CTTCTTCACTAAGTGC 1206

RESULT 15

US-08-486-756A-44/C  
 ; Sequence 44: Application US/08486756A  
 ; Patent No. 5981711

GENERAL INFORMATION:

APPLICANT: Zavada, Jan  
 APPLICANT: Pastorekova, Silvia  
 APPLICANT: Pastorek, Jeromir  
 TITLE OF INVENTION: MN Gene and Protein  
 NUMBER OF SEQUENCES: 86  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Leona L. Lauder  
 STREET: 6 Mariposa Court  
 CITY: Tiburon  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94920

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/486.756A  
 FILING DATE: 07-JUN-1995

CLASSIFICATION:

424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/260,190  
 FILING DATE: 15-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.  
 REGISTRATION NUMBER: 30,863  
 REFERENCE/DOCKET NUMBER: D-0021.3C  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-435-2034  
 TELEFAX: 415-435-0727

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 1334 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 DESCRIPTION: 6th MN intron  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-486-756A-44

Query Match 1.3%; Score 16; DB 2; Length 1334;  
 Best Local Similarity 100.0%; Pred. No. le-02;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 cttcactaagtgc 33

|||||

Db 1221 CTTCTTCACTAAGTGC 1206

Search completed: September 15, 2001, 11:34:11  
 Job time: 5493 sec